

**Vergleichende Fourier-Analyse von Bildkategorien  
mit hohem und niedrigem künstlerischen Anspruch**

**DISSERTATION**

**Zu Erlangung des akademischen Grades  
Doctor medicinae (Dr. med.)**

**Dem Rat der Medizinischen Fakultät der Friedrich Schiller Universität Jena  
vorgelegt von:**

**Tamara Melmer  
geboren am 18.02.1984 in Rum/Österreich**

Die vorliegende Dissertation wurde im Rahmen der Grundlagenforschung am Institut für Anatomie der Friedrich-Schiller-Universität Jena unter der wissenschaftlichen Aufsicht von Herrn Prof. Dr. Dr. Christoph Redies angefertigt. Sie beruht auf der folgenden Originalpublikation:

Melmer T, Amirshahi SA, Koch M, Denzler J, Redies C. 2013. From regular text to artistic writing and artworks: Fourier statistics of images with low and high aesthetic appeal. *Frontiers in Human Neuroscience*, 7: 106

**Gutachter:**

1. Prof. Dr. Dr. Christoph Redies, Jena
2. Prof. Dr. Joachim Denzler, Jena
3. PD Dr. Pascal Nicklas, Mainz

Datum der öffentlichen Verteidigung: 10.06.2015

## INHALTSVERZEICHNIS

ZUSAMMENFASSUNG.....	1
EINLEITUNG.....	2
1. Was ist Sehen?.....	2
1.1. Neuronale Prozesse der Netzhaut.....	3
1.2. Neuronale Prozesse des primären visuellen Cortex.....	4
2. Was ist Betrachten? .....	6
2.1. <i>Bottom-up</i> -Konstruktion des visuellen Bildes .....	6
2.2. <i>Top-down</i> -Konstruktion des visuellen Bildes .....	7
3. Was ist Lesen? .....	8
4. Zentrale Fragestellungen .....	10
4.1. Fourier-Analyse im Dienste funktionaler Fragestellungen .....	11
4.2. Fourier-Analyse im Dienste ästhetischer Fragestellungen .....	12
4.3. Fourier-Analyse im Dienste erweiterter ästhetischer Fragestellungen.....	12
ORIGINALPUBLIKATION: Melmer, Amirshahi, Koch, Denzler und Redies. 2013. From regular text to artistic writing and artworks: Fourier statistics of images with low and high aesthetic appeal. <i>Frontiers in Human Neuroscience</i> , 7: 106. ....	15
DISKUSSION.....	30
1. Kunst und ästhetische Erfahrung: Interdisziplinäre Definitionen und Voraussetzungen.....	31
2. Kunst versus Schrift: Alltägliche Druckschrift ist nicht skalierungsinvariant.....	36
3. Steigungsdifferenz und Isotropie: Korrelate künstlerischen Anspruchs.....	40
4. Ästhetische Erfahrung: Kein Anhalt für kulturellen Einfluss.....	42
LITERATURVERZEICHNIS.....	44
LEBENS LAUF.....	50
EHRENWÖRTLICHE ERKLÄRUNG.....	51

## ZUSAMMENFASSUNG

Das Forschungsgebiet der Neuroästhetik untersucht jene zerebralen Prozesse, die der ästhetischen Erfahrung zugrunde liegen, sowie jene Sinnesreize, durch welche diese evoziert wird. In der vorliegenden Arbeit wurde der Versuch unternommen, mittels Fourier-Analyse die strukturalen Eigenschaften von Schriftsystemen zu identifizieren, die für das menschliche Gehirn ‚interessant‘ bzw. im ästhetischen Sinne ‚fremd‘ erscheinen.

Frühere Studien (Redies et al, 2007a,b) haben gezeigt, dass sich das Fourier-Spektrum von Kunstbildern und anderen Bildkategorien mit ästhetischem Anspruch unabhängig von ihrem kulturellen oder zeitlichen Kontext durch Fraktalität und Isotropie auszeichnet. Dieses Ergebnis gab den Anstoß, Schriftkategorien mit unterschiedlichen ästhetischen Ansprüchen derselben Analyse zu unterziehen und zu erforschen, ob auch Schrift unabhängig von kulturellen und zeitlichen Einflüssen eine skalierungsinvariante Komposition und eine isotrope Signalverteilung im Fourier-Raum aufweist.

Die Ergebnisse der Experimente ergeben, dass die Parameter der Isotropie und die Steigungsdifferenz des Signals im Fourier-Raum mit dem künstlerischen Anspruch der jeweiligen Bildkategorie korrelieren. Der kulturelle und zeitliche Kontext der Schrift- und Kunstbilder zeigte keinen Einfluss auf das Ergebnis. Für alltägliche Druckschrift ohne künstlerischen Anspruch konnten keine skalierungs-invarianten Strukturen nachgewiesen werden.

Somit leistet die vorliegende Arbeit einen Beitrag zur Erforschung jener visuellen Stimuli, die im Betrachter eine ästhetische Erfahrung auslösen können. Die Erlangung tieferer Einblicke in die neuropsychologischen Prozesse der ästhetischen Erfahrung bleibt zukünftigen Studien vorbehalten.



## **EINLEITUNG**

Was ist Kunst? Obwohl diese Frage auf den ersten Blick trivial erscheint, stellt sich bei näherer Betrachtung schnell heraus, dass es sich hierbei um einen überaus komplexen Sachverhalt handelt. Bislang haben verschiedene geisteswissenschaftliche Schulen den Versuch unternommen, auf Basis psychologischer und kulturwissenschaftlicher Theorien eine allgemein gültige Definition des Begriffs ‚Kunst‘ zu erstellen (siehe z.B. Read, 1954; Hochberg, 1978; Jakobson, 1992). In den letzten Jahrzehnten hat sich außerdem das Forschungsgebiet der Experimentellen Ästhetik als naturwissenschaftliche Disziplin mit dieser Frage auseinandergesetzt. In eben diese Tradition stellt sich die vorliegende Arbeit: Mittels der Fourier-Analyse wurden die statistischen Eigenschaften von Kunstbildern und Schriftsystemen mit und ohne künstlerischem Anspruch verglichen, um Gemeinsamkeiten und Unterschiede festzustellen. Selbst wenn die Ergebnisse dieser Studie lediglich einen kleinen Beitrag zur Klärung der Definition von Kunst leisten können, geht die vorliegende Arbeit davon aus, dass jedem Kunstbegriff das Auslösen einer ästhetischen Erfahrung zugrunde liegen muss. In diesem Zusammenhang sollen im ersten Abschnitt die Rolle des Sehens als physiologische Leistung sowie die Rolle des Betrachtens als kognitive Leistung besprochen werden, zumal diese beiden Vorgänge die Grundlage der Auslösung einer ästhetischen Erfahrung darstellen.

### **1. Was ist Sehen?**

Es fällt mir schwer zu glauben, dass die Beziehung zwischen der Physiologie der Sehrinde und den Schöpfungen von Künstlern eine rein zufällige ist.

Semir Zeki (2011)

Wie einleitend erwähnt, will die vorliegende Arbeit jene Strukturen untersuchen, die visuelle Stimuli für das Gehirn im positiven Sinne ‚verfremden‘ (Jakobson, 1992), um so eine visuell getriggerte ästhetische Erfahrung zu erzeugen. Grundvoraussetzung dafür ist das Verständnis der neuronalen Prozesse, die ablaufen, sobald Licht auf die Netzhaut trifft.

## 1.1. Neuronale Prozesse der Netzhaut

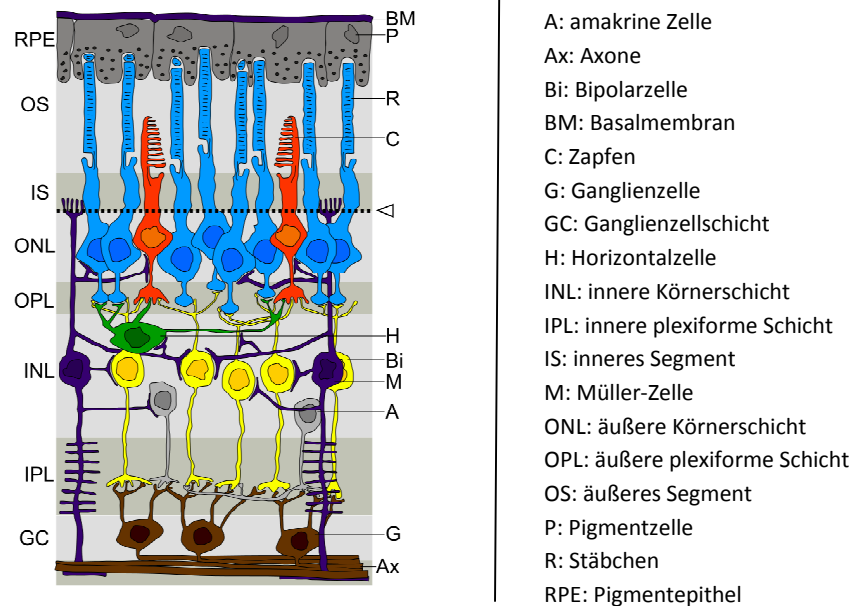


Abb. 1 - Schema der Netzhaut der Säugetiere (modifiziert nach Goldstein, 2002)

Treffen visuelle Stimuli auf die Netzhaut, werden durch diese spezialisierte Nervenzellen erregt. Bestimmte Komponenten des Seheindrucks werden extrahiert und in ein neuronales Signal umgewandelt, aus dem höhere Zentren des Gehirns schließlich einen Sinneseindruck rekonstruieren. Besonders die Arbeiten von Stephen Kuffler lieferten in den 1950er Jahren wesentliche Erkenntnisse über diese Verarbeitungsstrategie entlang der Sehbahn. Kuffler (1966) zeigte, dass die Kommunikation zwischen Nervenzellen auf zwei Prinzipien basiert: Neurone können einerseits einen exzitatorischen Effekt auf die Nachbarneurone haben, mit denen sie über Synapsen kommunizieren; in diesem Fall wird die Zahl der Aktionspotential der Nachbarneurone erhöht. Andererseits können Neurone inhibitorische Effekte ausüben, wodurch die Anzahl der Aktionspotenziale kommunizierender Neurone erniedrigt wird (Kandel, 2012). Die Gesamtheit exzitatorischer und inhibitorischer Signale entscheidet darüber, ob ein Signal an nachgeschaltete Zentren weitergeleitet wird oder nicht. Der britische Nobelpreisträger Charles Sherrington definierte diesen Prozess als *integrative Tätigkeit* (Levine, 2007) des Nervensystems. Kufflers Schüler David Hubel und Torsten Wiesel vertieften diese Studien zur neuronalen Kommunikation (Hubel und Wiesel, 1959) und zeigten im Tierexperiment, dass die Ganglienzellen der Netzhaut die erste Relaisstation der Sehbahn darstellen, die ihre Signale über die bipolaren Retinazellen von den auf Farbe spezialisierten Zapfen und den auf Helligkeit spezialisierten Stäbchen empfangen. Auch die

Entdeckung, dass Ganglienzellen selbst in Abwesenheit visueller Stimuli ununterbrochen Aktionspotentiale erzeugen, geht auf sie zurück. Kuffler zeigte außerdem, dass man die Frequenz dieser Aktionspotentiale am effizientesten durch kleine Lichtpunkte erhöhen kann, die auf jenes Territorium der Netzhaut gerichtet sind, aus dem die jeweilige Ganglienzelle ihre Eingänge bezieht – ihr sogenanntes *rezeptives Feld*. Dabei korrelieren die Frequenz der Aktionspotentiale mit der Intensität des Lichtreizes und die Dauer der Potentialblöcke mit der Dauer des dargebotenen Reizes. Die dritte Entdeckung, die Kuffler machte, verdeutlicht die Rolle des Kontrasts als grundlegendem Mechanismus des Sehsystems: Die Ganglienzellen der Netzhaut lassen sich in zwei Gruppen unterteilen. Die so- genannten *On-Zellen*, die vermehrt Aktionspotentiale abgeben, wenn Licht auf das Zentrum ihres rezeptiven Feldes trifft. Trifft Licht auf die Peripherie des rezeptiven Feldes, verringert sich die Frequenz der Aktionspotentiale. Die Off-Zellen reagieren umgekehrt und feuern vermehrt, wenn Licht auf die Peripherie ihres rezeptiven Feldes trifft und sie werden gehemmt, wenn ihr Zentrum stimuliert wird. Damit konnte Kuffler zeigen, dass das Sehsystem von Säugetieren nicht primär auf Lichtintensität reagiert, sondern auf Kontraste. Das System der *Zentrum-Umfeld-Hemmung* ist darauf ausgerichtet, Kontraständerungen effizient zu erfassen und an höhere Zentren weiterzuleiten, während sich nicht verändernde Muster als redundant erachtet und eliminiert werden.

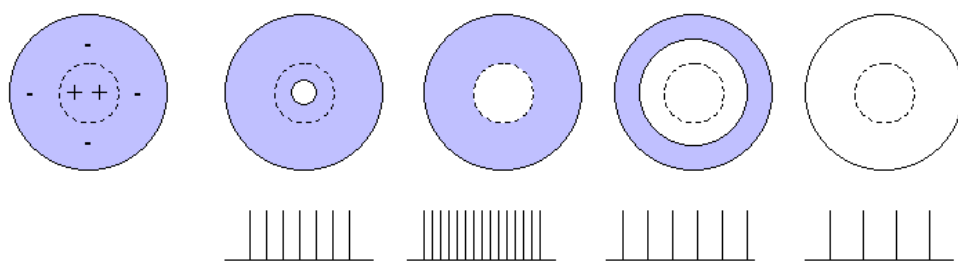


Abb. 2 - Aktionspotentiale einer On-Zelle (modifiziert nach Goldstein, 2002)

## 1.2. Neuronale Prozesse des primären visuellen Cortex

Wie das vorhergehende Kapitel gezeigt hat, ist die Netzhaut darauf spezialisiert, auf Basis der Kontraständerung zu entscheiden, ob visuelle Stimuli an nachgeschaltete Zentren weitergeleitet werden, oder nicht. Trotzdem enthält der visuelle Code der Netzhaut noch viel Information, die für die Rekonstruktion zentraler Formen wie Gesichter oder Szenen redundant ist, zumal dazu lediglich die wesentlichen Merkmale der Sehinformation benötigt

werden. Die Linie ist im Sinne einer durch Kontrast erzeugten Kante die wohl grundlegendste dieser Konstanten und kann daher vom Gehirn entsprechend effizient verarbeitet werden.

Als wegweisend für das Verständnis der Extraktion wesentlicher Merkmale in den der Netzhaut nachgeschalteten Zentren erwiesen sich die Arbeiten von David Hubel und Torsten Wiesel. Sie kamen zu der Erkenntnis, dass auch das Corpus geniculatum laterale des Thalamus über On- und Off-Zellen verfügt, die ihre Stimuli aus bestimmten Regionen der Netzhaut erhalten. Dasselbe gilt für hoch spezialisierte Zellgruppen im primär visuellen Cortex, in denen die Information bestimmter Netzhautareale verarbeitet wird (Hubel und Wiesel, 1959). Der Code des Corpus geniculatum laterale wird dabei nicht einfach übernommen. Vielmehr werden bestimmte lineare Komponenten extrahiert, während andere Informationen als redundant eliminiert werden.

Die Extraktion wesentlicher Merkmale geschieht zum ersten durch die sogenannten *einfachen Zellen*, deren rezeptive Feldern balkenförmig angeordnet sind. Hubel und Wiesel (1962) gingen davon aus, dass die längliche Anordnung durch die teilweise Überlappung konzentrischer rezeptiver Felder entsteht, wie aus Abbildung 3c hervorgeht. Einfache Zellen reagieren nicht nur auf Kanten, sondern zusätzlich auf deren räumliche Orientierung. Es gibt also Gruppen einfacher Zellen für senkrechte, waagrechte und schräge Linien. Der Seheindruck einer senkrechten Linie erhöht beispielsweise die Frequenz der Aktionspotenziale der auf senkrechte Orientierungen spezialisierten Gruppe einfacher Zellen. Zum zweiten werden wesentliche Bildinformationen durch *komplexe Zellen* extrahiert, deren rezeptive Felder konzentrisch organisiert sind. Hierbei gibt es Zellen, die sich auf eine bestimmte Bewegungsrichtung spezialisiert haben. Daneben existieren zum dritten die *hyperkomplexen Zellen*, die sensibel auf bestimmte Bewegungsrichtungen sowie auf Streifen und bestimmte Winkel reagieren (Hubel und Wiesel, 1962).

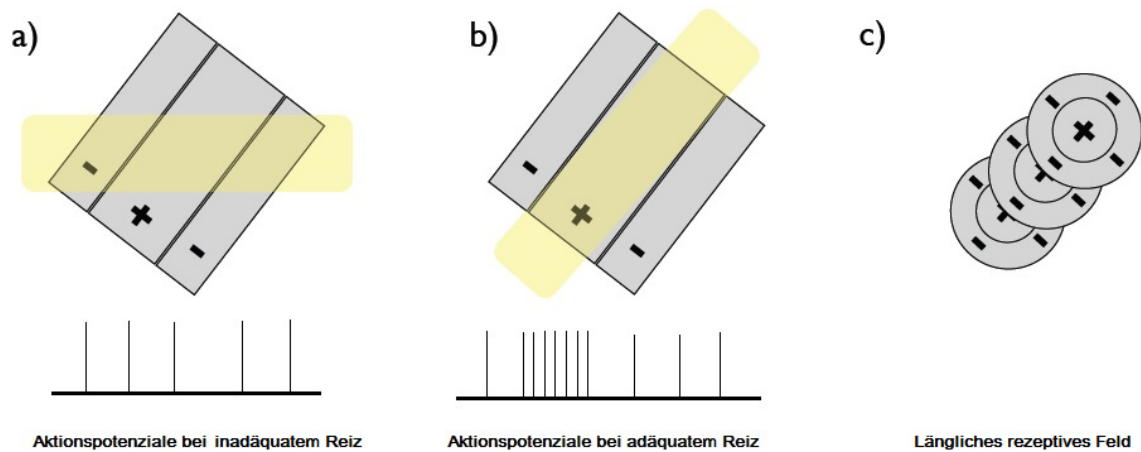


Abb. 3 - Verhalten und Eigenschaften einfacher Zellen (modifiziert nach Goldstein, 2002)

## 2. Was ist Betrachten?

Das vorhergehende Kapitel hat zwar die Prozesse der Merkmalerkennung erläutert, es konnte jedoch keine Antwort auf die Frage liefern, wie aus diesen einzelnen Komponenten ein in sich geschlossenes Bild rekonstruiert wird. Um diesen Vorgang nachvollziehbar zu machen, sollen in diesem Kapitel die Leistungen der Assoziationsareale des Gehirns besprochen werden. Zu diesem Zweck ist es sinnvoll, die Arbeiten des britischen Hirnforschers David Marr zu erwähnen, der auf Basis der Erkenntnisse Hubels und Wiesel eine bis heute gültige Theorie des Sehvorgangs entwickelte. Neben physiologischen Aspekten berücksichtigt er auch die Kognitionspsychologie.

### 2.1. Bottom-up-Konstruktion des visuellen Bildes

Wie Marr 1982 in seinem Werk *Vision* dargelegt, definiert er die bereits beschriebenen Vorgänge der Extraktion von Merkmalen als neurologische Vorgänge der *unteren Ebene*, die sich in der Netzhaut abspielen und dazu dienen, ein Objekt räumlich einzuordnen und dessen Farben zu erkennen. Zu den Prozessen der *mittleren Ebene*, die sich im primären visuellen Cortex abspielen, zählt die so genannte Konturintegration, deren Leistung darin besteht, die Umrisse eines Objekts zu erstellen. Im Sinne der Oberflächensegmentierung wird die geschlossene Form des Objekts außerdem von ihrem Hintergrund abgehoben.

Die Verarbeitung visueller Informationen auf der unteren und mittleren Ebene dient im Sinne eines *Bottom-up*-Vorgangs also der Selektion jener Merkmale, die sich effizient zu

einem geschlossenen Bild zusammensetzen lassen. Zu den Selektionskriterien, die dabei zum Einsatz kommen, zählen verschiedene Gesetzmäßigkeiten. So werden bevorzugt jene Farben, Größen und Richtungen zu einer Gruppe kombiniert, die einander räumlich *nahe* stehen und die das Gehirn im Sinne des Gestaltgesetzes ggf. selbst *fortsetzen* bzw. vervollständigen kann (Marr, 1982).

## **2.2. Top-down-Konstruktion des visuellen Bildes**

Die Verarbeitungsprozesse der *oberen Ebene* ermöglichen schließlich das Erkennen des visuellen Bildes im Sinne eines Gegenstandes, eines Körperteils oder einer Szene. Zu diesem Zweck wird die visuelle Information mit bereits gespeicherten Inhalten aus der Erinnerung abgeglichen und der entsprechenden Kategorie zugeordnet. Dieser Vorgang lässt sich in zwei Teilprozesse gliedern, nämlich den Abgleich entlang der *Was-Bahn* und den Abgleich entlang der *Wo-Bahn* (Goodale und Milner, 1992).

Die Neurone der Was-Bahn vernetzen den primär visuellen Cortex mit dem unteren Temporallappen. Sie beziehen ihren Input im Wesentlichen von den Zapfen der Netzhaut, sie ermöglicht es uns also, unseren Seheindruck farblich einzuordnen sowie zu bestimmen, ob es sich zum Beispiel um ein Objekt, eine Szene oder eine Person handelt. Die visuelle Information wird entlang der Was-Bahn also mit einem bereits bestehenden *Farben-* und *Formensystem* abgeglichen (Milner und Goodale, 2008).

Auf eben diesem Prinzip beruht nach Auffassung der amerikanischen Neurobiologin Margaret Livingstone die Wahrnehmung von Kunst durch das menschliche Sehsystem: Bei der Betrachtung von Umrissen bzw. Linien ist das Gehirn stets bemüht, das zu Sehende mit etwas bereits Erfahrenem zuzuordnen. Es ist also bestrebt, die zweidimensionale Bildinformation mit Erfahrungen im dreidimensionalen Raum im Einklang zu bringen. Es konstruiert daher Komponenten wie Höhe und Breite, obwohl derartige Informationen im Originalbild nur angedeutet sind (Conway und Livingstone, 2007).

Im Gegensatz zur Was-Bahn kann die Wo-Bahn keine Farben erkennen. Die zugehörigen Neurone vernetzen die primäre Sehrinde mit dem hinteren parietalen Cortex. Ihre Spezialisierung liegt in der Erkennung von Bewegung, indem sie die Adjustierung unseres Blicks durch Sakkaden unterstützt: fixierte Objekte werden durch schnelle Nachstellbewegungen des Auges stets im Bereich der Fovea centralis gehalten, dem Bereich der Netzhaut des schärfsten Sehens. Die Wo-Bahn reagiert sehr schnell auf Kontraständerungen

und kann Bewegungen dadurch frühzeitig und effizient nachvollziehen. Um die kognitiven Leistungen dieser beiden Bahnen miteinander in Einklang zu bringen, bedarf es der bewussten Aufmerksamkeit des Individuums (Pritchard et al, 1960). Das Prinzip der selektiven Aufmerksamkeit ist also ausschlaggebend für die Effizienz unseres Sehsystems und verdeutlicht einmal mehr die Rolle des *Verfremdungseffekts* (Jakobson, 1992) bzw. der *inherent interestingsness* (Hochberg, 1978) bei der Rezeption von Kunst: Erst der aus der Verfremdung resultierende, ästhetische Kontrast der visuellen Information erregt effizient die Aufmerksamkeit des Betrachters.

### **3. Was ist Lesen?**

Der Prozess des Lesens unterscheidet sich maßgeblich von der Betrachtung eines Kunstwerkes, zumal es sich beim Lesen von Text um die Erschließung der Semantik eines linearen Kodes handelt: Gelesen wird im westlichen Kulturraum Zeile für Zeile und von oben nach unten. Dem gegenüber existieren für das Betrachten eines Kunstwerkes keine einheitlichen Vorgaben. Wie Studien zu Augenbewegungen zeigten (Wooding et al, 2002; Quiroga und Pedreira, 2011), schweift der Blick des Rezipienten entlang eines individuellen Pfades über die Bildfläche eines Kunstwerks, dessen globale Struktur als flächiger Kode funktioniert. Es ist anzunehmen, dass sich diese verschiedenen Rezeptionsstrategien aus den spezifischen statistischen Eigenschaften der jeweiligen Bildkategorien ergeben. So zeichnet sich beispielsweise das Fourier-Spektrum bestimmter Kunstbilder durch wenig Anisotropie aus (Koch et al, 2010), grobe und feine Strukturen verteilen sich also gleichmäßig über alle Orientierungen. Dem gegenüber erweist sich Text als hoch anisotrop, was sich auf die klare räumliche Struktur der Zeilen zurückführen lässt. Aus diesem Grund wurden die räumlichen Orientierungen von Textbildern in der vorliegenden Arbeit getrennt betrachtet. Anders als bei der Analyse natürlicher Szenen, bei der das Fourier-Signal radiär über alle Frequenzen gemittelt wurde (siehe oben), ergibt sich die 1d-Spektren in der vorliegenden Arbeit aus der getrennten Betrachtung der kardinalen (horizontal und vertikal) und schrägen Orientierungen.

Um die spezifischen statistischen Unterschiede zwischen Text- und Kunstbildern zu bestimmen, wurden Bildkategorien herangezogen, die einen Übergang von Gebrauchstext zu Kunst mit ästhetischem Anspruch darstellen. Untersucht wurden Textbilder mit künstlerischem oder ästhetischem Anspruch wie ornamentale Schrift und Kalligrafie sowie

schriftähnliche Bilder mit ästhetischem Anspruch, deren Bildelemente mehr fragmentarisch nebeneinander stehen, als sich in ein homogenes Ganzes zu fügen, wie dies bei ornamentaler Kunst und den Werken des Abstrakten Expressionismus der Fall ist.

Zu beachten ist, dass im Zuge der Forschungen zur experimentellen Ästhetik bislang kein objektives Maß für den ästhetischen oder künstlerischen Gehalt eines Bildes entwickelt werden konnte. Die Verwendung der Begriffe ‚Kunst‘ und ‚Ästhetik‘ orientiert sich in der vorliegenden Arbeit daher an der Einschätzung der breiten Masse und mag Einzelpersonen mitunter ungerechtfertigt erscheinen. So argumentieren Typografen und Schriftsetzer auf der einen Seite sehr wohl für einen ästhetischen Anspruch von Gebrauchsschrift, während die breite Öffentlichkeit entsprechende Schriften nicht als künstlerisch wahrnimmt. Auf der anderen Seite stellen viele Einzelpersonen den ästhetischen Gehalt anerkannter Kunstbilder wie beispielsweise der Klecksbilder eines Jackson Pollock in Frage. Diese terminologische Unsicherheit reflektiert in letzter Instanz die variierenden Definitionen von Kunst aus philosophischer und kunsthistorischer Sicht. Entsprechend vorsichtig wird die vorliegende Arbeit mit diesen Begriffen umgehen. Darüber hinaus stellt sich in Anbetracht der enormen Vielzahl von Schriften verschiedener Kulturkreise die Frage, ob die statistischen Eigenschaften, die in der vorliegenden Studie beschrieben werden, tatsächlich der Adaption an die neuronale Kodierung der menschlichen Wahrnehmung genügen sollen (Changizi und Shimojo, 2005) und damit universaler Natur sind oder aber interkulturellen Unterschieden geschuldet sind. Aus diesem Grund wurden Bildkategorien aus dem westlichen, fernöstlichen und arabischen Kulturraum einbezogen.

Der Vorgang des Lesens basiert aus physiologischer Sicht im Wesentlichen auf der bereits erwähnten *Fixation* visueller Stimuli und erfordert daher die neuronalen Leistungen der höheren Ebene entlang der Wo-Bahn. Geübte Leser erkennen nicht einzelne Buchstaben, sondern ganze Wörter. Die visuelle Worterkennung beruht darauf, dass die einzelnen Wörter für eine Zeitspanne von jeweils etwa 0,3 bis 0,8 Sekunden (Martinez-Conde, Macknik und Hubel, 2004) im Bereich der Fovea centralis gehalten werden, ehe sich das Auge im Sinne einer Sakkade zum nächsten Fixationspunkt bewegt. Je nach Geübtheit des Lesers sind drei bis vier Fixationen pro Sekunde möglich. So schafft ein Schulkind, etwa 100 Wörter pro Minute zu lesen, weil ein Wort mitunter nicht als Ganzes erfasst wird, sondern durch mehrere Fixationspunkte erschlossen werden muss. Geübte Erwachsene erreichen drei bis



vier Fixationen pro Sekunde und lesen daher 180 bis 200 Wörter pro Minute (Martinez-Conde, Macknik und Hubel, 2004).

Der Vorgang des *fovealen Sehens* vollzieht sich in einem Abstand von bis zu 2,5 Grad Sehwinkel von der Fovea centralis (Westheimer und McKnee, 1975). Aufgrund ihrer Dichte an Nervenzellen sowie aufgrund der Geschwindigkeit der Sakkaden liefert foveales Sehen drei bis vier hoch auflösende Bilder pro Sekunde. Dem gegenüber liefert das sogenannte periphere Sehen, welches in einem Abstand von zwei bis neunzig Grad Sehwinkel von der Fovea passiert, bis zu neunzig Bilder pro Sekunde; deren räumliche Auflösung ist jedoch wesentlich geringer. Die im vorigen Kapitel beschriebenen Assoziationsareale kombinieren diese Teilbilder schließlich zu einem zusammenhängenden Textbild.

Von *Crowding* spricht man darüber hinaus, wenn die Bildinformation des fovealen und peripheren Sehens durch die Neurone der höheren Ebene nicht eindeutig voneinander getrennt werden kann. Eine Theorie sieht die Ursache dieses Effekts darin, dass sich Konturen, die sich näher als 0.5 Grad Sehwinkel (Bouma, 1970) nebeneinander befinden, nicht getrennt räumlich aufgelöst werden können. Die Frage nach dem idealen Buchstabenabstand ist deshalb zentral für die Gestaltung von Texten, um eine optimale Lesegeschwindigkeit zu erlauben.

Für die vorliegende Arbeit sind die Aspekte der Erkennbarkeit von Buchstaben bzw. Textbildern jedoch weniger relevant, als die Frage nach deren möglicherweise inhärenten künstlerischen Anspruch. Die Ausführungen der beigefügten Publikation sowie der darauf folgenden Diskussion versuchen demnach der Frage nachzugehen, ob Schrift als effizient prozessierbarer, visueller Stimulus ähnliche strukturelle Eigenschaften aufweist wie Kunstbilder (siehe unten).

#### **4. Zentrale Fragestellungen**

Das Forschungsgebiet der Neuroästhetik untersucht jene zerebralen Prozesse, die der ästhetischen Erfahrung zugrunde liegen, sowie jene Sinnesreize, durch welche diese Erfahrung hervorgerufen wird (Zeki, 2001). In der vorliegenden Arbeit wurde der Versuch unternommen, mittels Fourier-Analyse derartige strukturelle Eigenschaften in Schriftsystemen zu untersuchen. Ziel der experimentellen Arbeit war es also, jene Strukturen zu identifizieren, die für das menschliche Gehirn ‚interessant‘ bzw. im ästhetischen Sinne ‚fremd‘ erscheinen. Obwohl es unzulässig wäre, das menschliche Sehsystem als einen

„Fourier-Transformator“ zu definieren, ermöglicht die Analyse von Bildinformation im Fourier-Raum dennoch die Quantifizierung jener den visuellen Stimuli inhärenten Strukturen, die durch das menschliche Sehsystem einerseits extrahiert und verarbeitet werden und die andererseits eine ästhetische Empfindung im Betrachter auslösen können. In diesem Sinne gibt das Fourier-Spektrum eines Bildes Aufschlüsse über dessen formale Komposition und damit über dessen „Struktur“ im Sinne der Prager Schule um Roman Jakobson (siehe Diskussion, Abschnitt 1).

#### **4.1. Fourier-Analyse im Dienste funktionaler Fragestellungen**

Wie die Ausführungen des vorhergehenden Kapitels zeigten, filtern und transformieren spezialisierte Gruppen von Nervenzellen des menschlichen Sehsystems visuelle Informationen, um sie für die in höheren Zentren des Gehirns aufzubereiten, wo schließlich eine Interpretation des realen Objektes ermöglicht wird. In den 1960er Jahren gingen Sehforscher davon aus, dass das menschliche Auge zu diesem Zweck Lichtreize gemäß des Fourier-Algorithmus verarbeitet (Campbell und Robson, 1968). Selbst wenn diese These aus heutiger Sicht nicht mehr haltbar ist, war sie wegbereitend für die Erkenntnis, dass die sinusoidalen Einzelkomponenten des Lichts von spezialisierten Zellen getrennt erfasst und weitergeleitet werden, wie in den physiologischen Betrachtungen des Sehvorgangs besprochen wurde. Campbell und Robson (1968) gingen des Weiteren davon aus, dass das Gehirn aus den Fourier-Komponenten seine Interpretation der Lichtverhältnisse, Farben und Kontraste des Originalbildes rekonstruiert.

Wie in Studien aus den 1970er Jahren gezeigt werden konnte (Maffei und Fiorentini, 1973), spielen die Parameter der Ortsfrequenz, der Lage im Raum sowie des Kontrasts eine Schlüsselrolle bei der neuronalen Verarbeitung visueller Stimuli. Auf die Verarbeitung eben diese Parameter haben sich die Nervenzellen des menschlichen Sehsystems im Laufe der Evolution spezialisiert. In der Wahrnehmungspsychologie bezeichnet die Ortsfrequenz die Anzahl der Kanten pro Grad Sehwinkel, wobei eine „Kante“ im Wesentlichen durch den Kontrast zwischen zwei Flächen unterschiedlicher Helligkeit entsteht, die direkt nebeneinander liegen. Unscharfe Bilder weisen eine niedrige Anzahl von Kanten pro Grad Sehwinkel auf und erscheinen daher flächig, während auf Bildern mit hoher Ortsfrequenz auch feine Strukturen erkennbar sind. Genau genommen basiert also die Linie, die einleitend

bereits als Urform künstlerischen Ausdrucks bezeichnet wurde, auf dem Kontrast zwischen aufeinandertreffenden Flächen.

#### **4.2. Fourier-Analyse im Dienste ästhetischer Fragestellungen**

Während die Fourier-Analyse zur Erforschung der Funktion des menschlichen Sehsystems bereits auf eine längere Tradition zurückblickt, hat man ästhetische Fragestellungen erst in jüngster Zeit aufgeworfen. Wie kürzlich gezeigt wurde, hat das Fourier-Spektrum visuell ansprechender Bilder wie zum Beispiel Werke bildender Kunst westlichen und fernöstlichen Ursprungs (Graham und Field, 2007; Redies et al, 2007b; Graham und Redies, 2010), aber auch *graphic novels* (Koch et al, 2010) folgende Eigenschaft: Mit zunehmender Frequenz gehorcht die Abnahme der räumlich gemittelten (1d) Signalstärke dem Potenzgesetz  $1/f^2$  und weist im doppelt logarithmischen Koordinatensystem eine Steigung von -2 auf. Diese Besonderheit lässt sich dahingehend interpretieren, dass bestimmte visuell ansprechende Bildkategorien skalierungsinvariante Strukturen aufweisen (Graham und Redies, 2010).

Als ästhetisch empfundene Bilder teilen diese Eigenschaft mit komplexen natürlichen Szenen (Burton und Moorhead, 1987; Field, 1987; Tolhurst et al, 1992). Entsprechend werden Bilder, deren Struktur von natürlichen Szenen abweicht, mitunter als irritierend wahrgenommen (Fernandez und Wilkins, 2008; Juricevic et al, 2010; O'Hare und Hibbard, 2011). Das Sehsystem von Säugetieren hat im Laufe der Evolution gelernt, das sensorische Signal komplexer natürlicher Szenen durch spezifische neuronale Verschaltungen sparsam zu kodieren und diesen ‚verdichteten‘ Kode effizient zu verarbeiten (Olshausen und Field, 1996; Parraga et al, 2000; Vinje und Gallant, 2000; Hoyer und Hyvärinen, 2002; Simoncelli, 2003). Darauf basiert die Annahme, dass Künstler ihre Werke intuitiv an die Kodierungsstrategie des menschlichen Sehsystems anpassen (Redies, 2007; Redies et al, 2007a; Graham und Redies, 2010).

#### **4.3. Fourier-Analyse im Dienste erweiterter ästhetischer Fragestellungen**

Frühere Studien (Redies et al, 2007; Graham und Redies, 2010) haben gezeigt, dass sich das Fourier-Spektrum von Kunstbildern und anderen Bildkategorien mit ästhetischem Anspruch unabhängig von ihrem kulturellen oder zeitlichen Kontext durch Skalierungsinvarianz und einen hohen Grad an Isotropie auszeichnet. Dieses Ergebnis gab den Anstoß, auch

Schriftkategorien mit unterschiedlichen ästhetischen Ansprüchen derselben Analyse zu unterziehen und zu erforschen, ob auch Schrift unabhängig von kulturellen und zeitlichen Einflüssen ein skalierungsinvariantes Ortsfrequenz-Spektrum und eine isotrope Signalverteilung im Fourier-Raum aufweist. Es galt also zu klären, ob auch andere optische Signale, die von Menschenhand im alltäglichen Kontext erstellt wurden, ebenfalls dem besagten Potenzgesetz  $1/f^2$  gehorchen. Zu diesen Signalen zählen visuell wahrnehmbare Muster, die ohne künstlerischen Anspruch erzeugt wurden, wie zum Beispiel Gebrauchstext. Im Zuge funktionseller Forschungsfragen wurden die räumlichen Eigenschaften von Buchstaben im Hinblick auf Lesbarkeit und Erkennbarkeit in den letzten Jahrzehnten intensiv erforscht (Pelli et al, 2006; Chung und Tjan, 2007; Tyler und Likova, 2007; Chung und Tjan, 2009). Dabei konzentrierte man sich vor allem auf die Aspekte der Ortsfrequenz und des Buchstabenabstandes als Faktoren für leichtere Erkennbarkeit, bessere Lesbarkeit und geringere Ermüdung beim Lesen (Solomon und Pelli, 1994; Majaj et al, 2002; Wilkins et al, 2007; Nandy und Tjan, 2008; Oruc und Landy, 2009; Jainta et al, 2010). Dagegen wurden die globalen statistischen Eigenschaften des Schriftbildes, die möglicherweise die ästhetische Funktion eines Schriftsystems bedingen können, bisher kaum beachtet (beispielsweise Wilkins, 1995). Tatsächlich handelt es sich bei der Lesbarkeit und der Ästhetik von Schrift um zwei eigenständige Kategorien, die mitunter nicht miteinander einhergehen: Künstlerisch gestaltete Schriften mit aufwändigen Ornamenten oder deformierten Buchstaben werden oft als ästhetisch empfunden, sind jedoch tendenziell schwer lesbar, wie sich am Beispiel der chinesischen Gasschrift zeigen lässt. Dem gegenüber ermöglicht die Typografie alltäglicher Gebrauchsschrift eine gute Lesbarkeit, während ästhetische Ansprüche zweitrangig erscheinen.

Die vorliegende Arbeit vergleicht erstmals die statistischen Eigenschaften alltäglicher Textbilder mit den Datensätzen einfarbiger Kunstbilder, die in früheren Studien bereits analysiert wurden (Redies et al, 2007b). Selbst wenn Gebrauchstexte nicht zwingend ästhetischen Ansprüchen genügen sollen, geht man davon aus, dass ihre Struktur der Arbeitsweise des menschlichen Sehsystems entgegenkommen könnte (Wilkins, 1995; Changizi und Shimojo, 2005; Changizi et al, 2006; Jainta et al, 2010). Aus diesem Grund untersucht diese Studie einzelne Buchstaben nicht auf statistische Eigenschaften und Lesbarkeit, sondern nimmt die globale Struktur von Textbildern in den Fokus. Als Vorbild diente dabei eine vorherige Studie zum Vergleich des Fourier-Spektrums zwischen

japanischer Grasschrift und alltäglicher Druckschrift (Ozawa, 1994). Im Sinne einer neuroästhetischen Herangehensweise leisten die Ergebnisse dieser Studie einen Beitrag zum Verständnis der Wechselwirkung zwischen den statistischen Eigenschaften von Text und den entsprechenden neuronalen Wahrnehmungsprozessen des menschlichen Sehsystems. Durch den Vergleich von Textbildern mit anderen von Menschen geschaffenen Bildern tragen die Resultate außerdem zum Verständnis jener strukturalen Eigenschaften bei, die Schrift für das menschliche Gehirn möglicherweise ‚interessant‘ erscheinen lassen könnten.

Zusammenfassend sollen also die folgenden spezifischen Fragen beantwortet werden:

- (1) Weist Schrift ähnlich wie Kunstbilder skalierungsinvariante Strukturen im Fourier-Raum auf?
- (2) Korrelieren Steigung und Isotropie des Fourier-Signals mit dem ästhetischen Anspruch der jeweiligen Bildkategorie?
- (3) Bestehen interkulturelle Unterschiede zwischen Bildkategorien mit ästhetischem Anspruch?



# From regular text to artistic writing and artworks: Fourier statistics of images with low and high aesthetic appeal

Tamara Melmer<sup>1</sup>, Seyed A. Amirshahi<sup>1,2</sup>, Michael Koch<sup>1,2</sup>, Joachim Denzler<sup>2</sup> and Christoph Redies<sup>1\*</sup>

<sup>1</sup> Experimental Aesthetics Group, Institute of Anatomy I, University of Jena School of Medicine, Jena University Hospital, Jena, Germany

<sup>2</sup> Computer Vision Group, Department of Computer Science, Friedrich Schiller University, Jena, Germany

## Edited by:

Mohamed L. Seghier, University College London, UK

## Reviewed by:

Shozo Tobimatsu, Kyushu University, Japan

Arnold J. Wilkins, University of Essex, UK

## \*Correspondence:

Christoph Redies, Experimental Aesthetics Group, Institute of Anatomy I, Jena University Hospital, Teichgraben 7, D-07743 Jena, Germany.  
e-mail: redies@mti.uni-jena.de

The spatial characteristics of letters and their influence on readability and letter identification have been intensely studied during the last decades. There have been few studies, however, on statistical image properties that reflect more global aspects of text, for example properties that may relate to its aesthetic appeal. It has been shown that natural scenes and a large variety of visual artworks possess a scale-invariant Fourier power spectrum that falls off linearly with increasing frequency in log-log plots. We asked whether images of text share this property. As expected, the Fourier spectrum of images of regular typed or handwritten text is highly anisotropic, i.e., the spectral image properties in vertical, horizontal, and oblique orientations differ. Moreover, the spatial frequency spectra of text images are not scale-invariant in any direction. The decline is shallower in the low-frequency part of the spectrum for text than for aesthetic artworks, whereas, in the high-frequency part, it is steeper. These results indicate that, in general, images of regular text contain less global structure (low spatial frequencies) relative to fine detail (high spatial frequencies) than images of aesthetics artworks. Moreover, we studied images of text with artistic claim (ornate print and calligraphy) and ornamental art. For some measures, these images assume average values intermediate between regular text and aesthetic artworks. Finally, to answer the question of whether the statistical properties measured by us are universal amongst humans or are subject to intercultural differences, we compared images from three different cultural backgrounds (Western, East Asian, and Arabic). Results for different categories (regular text, aesthetic writing, ornamental art, and fine art) were similar across cultures.

**Keywords:** writing systems, calligraphy, ornamental art, abstract artworks, spatial frequency, scale invariance, experimental aesthetics

## INTRODUCTION

During the last decades, the spatial characteristics of letters and their influence on readability and letter identification have been studied in considerable detail (Pelli et al., 2006; Chung and Tjan, 2007; Tyler and Likova, 2007; Chung and Tjan, 2009). A particular focus has been on the spatial frequency components and spacing requirements that facilitate letter identification and improve readability and reading comfort (Solomon and Pelli, 1994; Majaj et al., 2002; Wilkins et al., 2007; Nandy and Tjan, 2008; Oruc and Landy, 2009; Jainta et al., 2010). The statistical image properties that relate to more global aspects of text images, for example properties that may relate to aesthetic aspects of writing systems, have received less attention (for an example, see Wilkins, 1995). Readability and aesthetics are two independent aspects of writing. Artistic writing with intricate ornaments or deformed letters may be highly aesthetic but it can sometimes be difficult to read, for example, Chinese cursive script. In contrast, ordinary (non-artistic) typographic writing is easy to read but may not necessarily be aesthetic.

It has recently been shown that a large variety of aesthetic and other visually pleasing images, including art images of Western and Eastern provenance (Graham and Field, 2007; Redies et al., 2007b; Graham and Redies, 2010) and graphic novels (Koch et al.,

2010), exhibit specific statistical properties in their Fourier spectra: with increasing frequency, radially averaged (1d) power falls off according to a power law with a slope of around  $-2$  in log-log plots, i.e., the power spectrum displays  $1/f^2$  characteristics. This property implies that subsets of visually pleasing images possess a scale-invariant structure in the Fourier domain. Aesthetic images share this property with complex natural scenes (Burton and Moorhead, 1987; Field, 1987; Tolhurst et al., 1992). Vice versa, images that deviate from natural scene statistics can induce visually discomfort (Fernandez and Wilkins, 2008; Juricevic et al., 2010; O'Hare and Hibbard, 2011). Because the mammalian visual system is adapted to process natural scenes with an efficient and sparse sensory code (Olshausen and Field, 1996; Parraga et al., 2000; Vinje and Gallant, 2000; Hoyer and Hyvärinen, 2002; Simoncelli, 2003), it has been proposed that artists create aesthetic images by adapting their artworks to this type of sensory coding in the human visual system (Redies, 2007; Redies et al., 2007a; Graham and Redies, 2010).

It remains unclear, however, whether other types of visual patterns that are produced by humans for viewing in everyday life, also possess  $1/f^2$  characteristics. Examples are visual patterns that are created without obvious aesthetic intent, such as regular text. In the present study, we therefore compared the statistical



properties of regular text with a set of monochrome graphic art analyzed previously (Redies et al., 2007b). Although regular text is not necessarily aesthetic, it may also be adapted or optimized to particular aspects of visual perception, as suggested previously (Wilkins, 1995; Changizi and Shimojo, 2005; Changizi et al., 2006; Jainta et al., 2010). It is therefore of interest to study the statistical properties of text images, not only with respect to local properties, such as the readability of individual letters and words (see above), but also in terms of the global appearance of text images, as suggested previously in a preliminary study on the Fourier spectrum of two examples of Japanese calligraphy and regular print (Ozawa, 1994).

The action of reading is likely to differ from viewing artworks because reading typographic text encompasses the deciphering of a linear code with semantic content (for example, reading line by line from left to right, or top to bottom). In contrast, viewing artworks is much less constrained and the composition of artworks allows the free exploration of global image structure, as shown by eye tracking studies (Wooding et al., 2002; Quiroga and Pedreira, 2011). The differences in viewing strategies likely correspond to differences in statistical image properties. For example, subsets of aesthetic art images were shown to have Fourier spectral properties that are rather uniformly distributed across image orientations (i.e., low anisotropy; Koch et al., 2010). In contrast, images of text are highly anisotropic due to the oriented structure of the lines of text. Moreover, the spatial frequency amplitude that corresponds to the distance between text lines can be expected to be high compared to other frequencies. In the present study of the Fourier spectra of text images, we therefore studied cardinal (horizontal and vertical) and oblique orientations separately to obtain 1d spectra, rather than radially averaging across all orientations together, as usually done when analyzing natural scenes and aesthetic images (see above).

To more closely define the differences in statistical properties between text images and aesthetic art images, we included image categories at the transition between regular text and aesthetic art. Specifically, we studied (1) images of artistic or aesthetic writing (ornamental writing, calligraphy), and (2) aesthetic images that are similar to text images in that they are composed of multiple, largely independent pictorial elements placed side-by-side (ornamental art and abstract expressionist art).

Note that in current experimental research on aesthetics, there is no universally accepted or independent measure for the degree to which an image is aesthetic or artistic. By using these terms for different categories of images in the present study, we rely on views that are commonly held by the general public. Our classifications may well be in conflict with the opinions of individual persons. For example, on the one hand, typographers might claim that the creation of typographic letters is a highly artistic endeavor whereas most of the general public will not regard images of regular printed text as artworks. On the other hand, individual viewers may consider highly acclaimed artworks, such as the drip paintings by Jackson Pollock, unaesthetic. This terminological uncertainty is reflected in the many views on art and aesthetics that abound in philosophy and art history. Consequently, the usage of these terms in the present study should be treated with caution.

Moreover, in view of the great variety of writing styles in different cultures, it was unclear whether any of the statistical properties measured in the present study are universal amongst humans (Changizi and Shimojo, 2005) or are subject to inter-cultural differences. We therefore compared images from three different cultural backgrounds (Western, East Asian, and Arabic).

Results from the present study contribute to the knowledge on the relationship between the statistical properties of text images and their perceptual processing by the human visual system. By comparing text images to other types of images produced by humans, the present findings contribute also to our understanding of what makes text images special to the human brain.

## MATERIALS AND METHODS

### IMAGE DATA

The data analyzed in this study include image databases of regular and ornate text of diverse cultures of writing, calligraphy, artistic ornaments of three cultural backgrounds (Western, Arabic, and East Asian), and East Asian and Western fine art (Table 1). A total of 1611 images were analyzed.

#### *Regular print, handwriting, ornate print, and calligraphy*

For the analysis of text images, we scanned the largest possible square section comprising eight lines of monochrome original text. Care was taken to select original print samples reproduced at a high quality and at a size that was sufficiently large. Scanning was performed at a high resolution (400 dpi) in 8-bit gray scale with a scanner (Perfection 3200 Photo, Epson, Nagano, Japan) that was calibrated as described previously (Redies et al., 2007a). Subsequently, resizing each image to  $1024 \times 1024$  pixels reduced the resolution. For each category of text, different original documents were used. The number of sections taken from one original document ranged from about 1 to 4.

For Latin serif and sans serif fonts, a sample text was set in 77 serif fonts and in 60 sans serif fonts with the Photoshop program (Adobe, Mountain View, CA). For international serif fonts, 119 examples of the same text from different writing systems were generated with the Photoshop program, including samples from Europe (Latin, Georgian, Cyrillic, Greek), the Middle East (Hebrew, Arabic), North America (Cherokee), North India (Devanagari, Gujarati, Nagari, Oriya), South India (Sinhala, Tamil, Telugu), South East Asia (Laotian, Khmer, Thai), Africa (Ethiopian), and the Far East (Chinese, Japanese, Korean). By the same method, examples of ornate print (117 Latin samples, 13 Chinese samples, 80 Arabic samples) were generated.

To assess differences within one font, we generated 30 samples each of different text passages that were set in Times New Roman font (Latin serif), Arial font (Latin sans serif), and a Georgian font (serif font), respectively. Moreover, we analyzed digital images directly for analysis without printing and scanning. As a control, we compared the digital images with the same images printed on paper and scanned as described above. The differences between the different types of images were small (data not shown).

For Latin and Carolingian handwriting, examples were scanned from two books on these subjects (Menz, 1912; Klemm, 1998). We also gathered 39 different examples of Arabic calligraphy by scanning three different books on the subject. The

**Table 1 | Slope values for art images and text images.**

	<i>n</i>	5–40 cycles per image		40–256 cycles per image		Slope difference
		Mean slope	Mean sigma	Mean slope	Mean sigma	
REGULAR PRINT						
Latin serif	77					
Horizontal		0.19 ± 0.20	0.062 ± 0.017	−3.62 ± 0.08	0.061 ± 0.015	3.81 ± 0.58
Oblique		−0.49 ± 0.28	0.061 ± 0.029	−3.44 ± 0.27	0.029 ± 0.011	2.95 ± 0.49
Vertical		−0.68 ± 0.57	0.182 ± 0.170	−2.91 ± 0.25	0.061 ± 0.028	2.23 ± 0.65
Latin sans serif	60					
Horizontal		0.09 ± 0.17	0.073 ± 0.016	−3.52 ± 0.04	0.093 ± 0.011	3.61 ± 0.74
Oblique		−0.52 ± 0.27	0.058 ± 0.022	−3.71 ± 0.11	0.045 ± 0.013	3.18 ± 0.62
Vertical		−0.63 ± 0.64	0.166 ± 0.163	−3.43 ± 0.13	0.070 ± 0.021	2.80 ± 0.63
International serif	119					
Horizontal		−0.31 ± 0.23	0.067 ± 0.016	−3.49 ± 0.07	0.064 ± 0.009	3.10 ± 0.65
Oblique		−0.78 ± 0.29	0.059 ± 0.023	−3.47 ± 0.19	0.036 ± 0.010	2.65 ± 0.61
Vertical		−0.71 ± 0.32	0.152 ± 0.127	−3.34 ± 0.20	0.063 ± 0.022	2.59 ± 0.68
HANDWRITING						
Carolingian	36					
Horizontal		−0.04 ± 0.19	0.058 ± 0.015	−3.72 ± 0.08	0.028 ± 0.005	3.69 ± 0.37
Oblique		−0.73 ± 0.27	0.051 ± 0.023	−3.41 ± 0.19	0.022 ± 0.006	2.67 ± 0.33
Vertical		−0.90 ± 0.30	0.187 ± 0.161	−3.87 ± 0.15	0.023 ± 0.013	2.97 ± 0.33
Latin handwriting	35					
Horizontal		−1.03 ± 0.21	0.041 ± 0.013	−3.48 ± 0.15	0.010 ± 0.001	2.45 ± 0.41
Oblique		−1.15 ± 0.32	0.045 ± 0.020	−3.45 ± 0.23	0.012 ± 0.005	2.30 ± 0.28
Vertical		−1.02 ± 0.33	0.129 ± 0.065	−3.53 ± 0.12	0.012 ± 0.004	2.51 ± 0.46
ORNATE PRINT						
Latin ornate print	132					
Horizontal		−0.67 ± 0.27	0.052 ± 0.017	−3.09 ± 0.16	0.035 ± 0.011	2.41 ± 0.71
Oblique		−0.81 ± 0.61	0.053 ± 0.022	−2.75 ± 0.35	0.028 ± 0.011	1.94 ± 0.67
Vertical		−0.82 ± 0.37	0.178 ± 0.141	−2.62 ± 0.33	0.034 ± 0.014	1.80 ± 0.74
Chinese ornate print	13					
Horizontal		−0.68 ± 0.35	0.090 ± 0.028	−3.21 ± 0.05	0.052 ± 0.005	2.53 ± 0.45
Oblique		−1.15 ± 0.20	0.080 ± 0.029	−2.94 ± 0.14	0.025 ± 0.006	1.79 ± 0.31
Vertical		−1.16 ± 0.21	0.147 ± 0.138	−2.69 ± 0.17	0.045 ± 0.007	1.53 ± 0.54
Arabic ornate print	80					
Horizontal		−0.79 ± 0.42	0.036 ± 0.011	−3.55 ± 0.06	0.113 ± 0.010	2.75 ± 0.28
Oblique		−1.06 ± 0.24	0.053 ± 0.024	−3.54 ± 0.12	0.044 ± 0.016	2.48 ± 0.16
Vertical		−0.44 ± 0.27	0.166 ± 0.107	−3.56 ± 0.08	0.147 ± 0.045	3.12 ± 0.32
CALLIGRAPHY						
Arabic cursive	39					
Horizontal		−0.91 ± 0.47	0.043 ± 0.017	−3.29 ± 0.14	0.043 ± 0.017	2.38 ± 0.92
Oblique		−0.86 ± 0.25	0.048 ± 0.021	−3.17 ± 0.17	0.030 ± 0.011	2.31 ± 0.81
Vertical		−0.88 ± 0.43	0.109 ± 0.059	−3.51 ± 0.15	0.035 ± 0.014	2.63 ± 0.90
Chinese cursive	92					
Horizontal		−0.99 ± 0.28	0.117 ± 0.031	−3.52 ± 0.06	0.034 ± 0.007	2.54 ± 0.63
Oblique		−0.95 ± 0.29	0.057 ± 0.027	−3.27 ± 0.13	0.023 ± 0.009	2.31 ± 0.43
Vertical		−1.25 ± 0.53	0.050 ± 0.016	−3.42 ± 0.12	0.034 ± 0.011	2.17 ± 0.53
ORNAMENTS						
Grotesque ornaments	69					
Horizontal		−1.08 ± 0.24	0.066 ± 0.017	−2.65 ± 0.06	0.016 ± 0.002	1.58 ± 0.55
Oblique		−1.02 ± 0.25	0.051 ± 0.020	−2.64 ± 0.07	0.013 ± 0.002	1.62 ± 0.43
Vertical		−1.27 ± 0.64	0.056 ± 0.021	−2.62 ± 0.06	0.016 ± 0.004	1.34 ± 0.51

(Continued)



Table 1 | Continued

	<i>n</i>	5–40 cycles per image		40–256 cycles per image		Slope difference
		Mean slope	Mean sigma	Mean slope	Mean sigma	
<b>Arabic ornaments</b>	253					
Horizontal		$-0.90 \pm 0.31$	$0.069 \pm 0.019$	$-2.89 \pm 0.09$	$0.017 \pm 0.004$	$2.04 \pm 0.86$
Oblique		$-0.79 \pm 0.30$	$0.070 \pm 0.026$	$-2.79 \pm 0.12$	$0.015 \pm 0.004$	$2.03 \pm 0.76$
Vertical		$-0.94 \pm 0.46$	$0.066 \pm 0.022$	$-2.83 \pm 0.11$	$0.017 \pm 0.005$	$1.94 \pm 0.74$
<b>East Asian ornaments</b>	78					
Horizontal		$-2.10 \pm 0.28$	$0.052 \pm 0.020$	$-3.62 \pm 0.13$	$0.003 \pm 0.002$	$1.55 \pm 0.70$
Oblique		$-2.02 \pm 0.24$	$0.053 \pm 0.020$	$-3.96 \pm 0.26$	$0.018 \pm 0.006$	$1.94 \pm 0.81$
Vertical		$-1.97 \pm 0.22$	$0.053 \pm 0.015$	$-3.76 \pm 0.11$	$0.012 \pm 0.003$	$1.78 \pm 0.82$
<b>FINE ART</b>						
<b>Abstract expressionism</b>	119					
Horizontal		$-1.80 \pm 0.24$	$0.047 \pm 0.02$	$-2.66 \pm 0.08$	$0.028 \pm 0.01$	$0.86 \pm 0.65$
Oblique		$-1.76 \pm 0.31$	$0.046 \pm 0.02$	$-2.51 \pm 0.11$	$0.019 \pm 0.00$	$0.74 \pm 0.72$
Vertical		$-1.84 \pm 0.29$	$0.045 \pm 0.02$	$-2.67 \pm 0.08$	$0.024 \pm 0.00$	$0.82 \pm 0.64$
<b>European fine art</b>	200					
Horizontal		$-2.13 \pm 0.24$	$0.045 \pm 0.01$	$-1.95 \pm 0.09$	$0.011 \pm 0.00$	$-0.18 \pm 0.81$
Oblique		$-2.20 \pm 0.30$	$0.044 \pm 0.02$	$-1.86 \pm 0.15$	$0.009 \pm 0.00$	$-0.34 \pm 0.81$
Vertical		$-2.15 \pm 0.28$	$0.043 \pm 0.02$	$-1.89 \pm 0.12$	$0.043 \pm 0.02$	$-0.25 \pm 0.73$
<b>East Asian fine art</b>	209					
Horizontal		$-2.05 \pm 0.25$	$0.046 \pm 0.01$	$-2.34 \pm 0.08$	$0.012 \pm 0.00$	$0.19 \pm 0.70$
Oblique		$-2.13 \pm 0.27$	$0.044 \pm 0.02$	$-2.20 \pm 0.14$	$0.011 \pm 0.00$	$0.07 \pm 0.68$
Vertical		$-2.13 \pm 0.26$	$0.042 \pm 0.01$	$-2.26 \pm 0.09$	$0.012 \pm 0.00$	$0.14 \pm 0.73$

background of the images was rendered white by subtracting it in the Photoshop program. As a control, we also analyzed the original scans of the same images. Results reveal differences in the low-frequency range, most likely caused by the paper structure. However, these differences have only a minor influence on the slopes (data not shown).

From reproductions in various textbooks on East Asian calligraphy, 92 examples of Chinese calligraphy representing different styles and periods were scanned. Some of the images showing eight lines of text were not square due to a smaller size of the original artworks in the other direction (fewer columns of text). Before their reduction to  $1024 \times 1024$  pixels, the images were padded according to square ones with the MatLab program by adding a uniform border with a gray level that was equal to the mean gray level of the scanned image, as described previously (Redies et al., 2007a).

### Ornaments

For Western grotesque ornaments, 69 different samples were scanned from a textbook (Warncke, 1979), as described above. With a digital camera (Canon, Ixus 400), 78 photographs of East Asian ornaments were taken from samples of 17th and 18th century Chinese and Japanese porcelain (vases, pots, and dishes) that were on display at the Dresden Porcelain Collection in Dresden, Germany. Square details from the objects were analyzed (for an example, see Figure 13A). Photographs of Arabic ornaments (253 images) were taken with a digital camera (Canon EOS 500D) from interior and exterior wall reliefs of the Alhambra Palace complex in Grenada, Spain. The palace represents an

example of 14th century Moorish architecture (for an example, see Figure 12A).

### Western and East Asian fine art and abstract expressionism

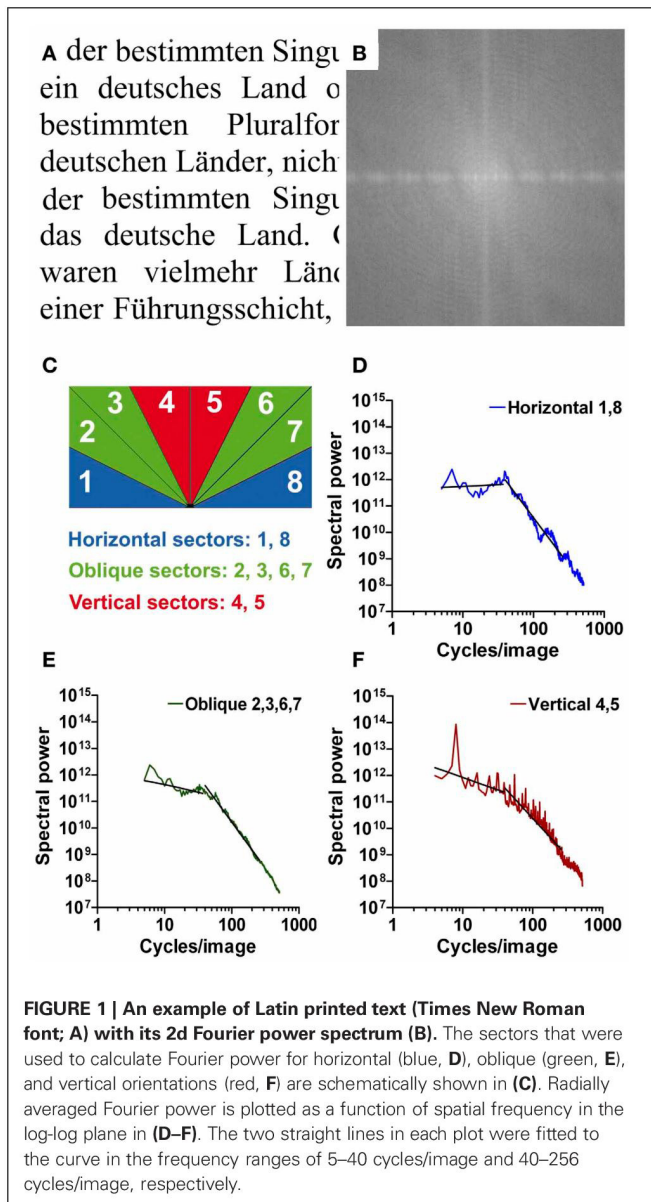
For Western art, a previously analyzed dataset of 200 examples of monochrome graphic art of Western provenance was used (Redies et al., 2007b). For East Asian (China, Japan, and Korea) fine art, 209 images of monochrome paintings were collected. Both datasets were scanned from diverse high-quality art books on the subject, as described above, and represented a large variety of graphic styles, subject matters, techniques, centuries, and artists. Largest possible square details from the artworks were analyzed. One-hundred nineteen examples of monochrome Abstract Expressionist art by four artists (32 paintings by Jackson Pollock, 18 paintings by Jean Dubuffet, 59 paintings by Cy Twombly, and 10 paintings by Christian Dotremont) were scanned from art-books. Images were padded according to square ones, as described above. To render the images more similar to those of printed text, the background was subtracted from the images. The calculated slopes did not differ substantially between the original scanned images and the images after background subtraction (data not shown). As another control, we compared largest possible details of the East Asian art images to padded versions of the same images. Again, differences were small (data not shown).

### IMAGE ANALYSIS

#### Radial averaging of Fourier power

Image analysis was performed using Matlab. If required, images were resized to a resolution of  $1024 \times 1024$  pixels by bicubic

interpolation using the *imresize* function in Matlab. For each image, the power spectrum (amplitude squared) was obtained by using an efficient algorithm for computing the discrete Fourier transform (2d Fast Fourier Transform). The 2d Fourier power spectrum of each image (for example, see **Figure 1B**) was divided into eight equal sectors (**Figure 1C**). For each sector, the 2d spectrum was transformed to a 1d power spectrum by rotational averaging for each frequency (Redies et al., 2007b). Values were combined for (near-)horizontal orientations (sector 1 [0–22.5°] and sector 8 [157.5–180°]; blue in **Figure 1C**), (near-)vertical orientations (sectors 4, 5 [67.5–112.5°]; red in **Figure 1C**), and oblique orientations (sectors 2, 3 [22.5–67.5°] and sectors 6, 7 [112.5–157.5°]; green in **Figure 1C**). Power was then plotted for horizontal orientations (**Figure 1D**), oblique orientations (**Figure 1E**), and vertical orientations (**Figure 1F**) as a function of spatial frequency in the log-log plane.



For regular print, the resulting plots consisted of two roughly linear parts: a low-frequency part (range 5–40 cycles per image, cpi) with a shallow slope and a high-frequency part (range 40–256 cpi) with a steeper slope. This finding was similar for horizontal, oblique, and vertical orientations. As expected with images of eight horizontal lines of text, a prominent peak at 8 cpi was observed for vertical orientations (**Figure 1F**).

### Slope of 1d Fourier plots

To measure the slopes of the curves in the two parts of the frequency spectrum, data points were binned at regular frequency intervals in the log-log plane and a least-squares fit of a line was performed separately for each of the two parts of the spectrum, as described previously (Redies et al., 2007b). Compared to the fitting to one continuous second-order polynomial function with three parameters, the fitting of two separate lines (with two intercepts and two slopes as parameters) allowed us to relate our present results more directly to previous slope measurements (Burton and Moorhead, 1987; Tolhurst et al., 1992; Graham and Field, 2007; Redies et al., 2007a,b; Koch et al., 2010).

For the high-frequency part, fitting was restricted to frequencies up to 256 cpi to minimize artifacts due to rectangular sampling and raster screen. For the low-frequency part, fitting was restricted to frequencies down to 5 cpi to avoid absent sample points for some orientations and to exclude information that is not of interest (artifacts due to uneven illumination and mean gray level). Moreover, for vertical orientations, values corresponding to the peak at around 8 cpi (7–9 cpi) were not included in the line fitting. As a measure of the goodness of the fit, we determined the mean deviation of the data points from the fitted lines (sigma in **Table 1**). **Table 1** lists the values as means for each image category [ $\pm 1$  standard deviation (SD)].

Because a characteristic difference between images of artworks and text seemed to be the change in the log-log plots of radially averaged Fourier power at around 40 cpi (see Results), we also calculated the difference between the slopes of the low-frequency and high-frequency parts.

### Anisotropy

As outlined in the Introduction section, the Fourier power spectrum of text images is likely to differ across orientations (anisotropy) because of the horizontal text lines (or vertical text lines in the case of Chinese writing). To analyze this anisotropy, we determined the average absolute difference between the power values for horizontal orientations (sectors 1 and 8 in **Figure 1C**) and vertical orientations (sectors 4 and 5 in **Figure 1C**) for each image. To calculate the difference, data were sampled at equal frequency intervals in the log-log plots of Fourier power for each image (see, e.g., **Figures 1D–F**). Differences were normalized to the mean power for vertical and horizontal orientations for each data point.

## RESULTS

**Figures 1, 2** illustrate the type of Fourier analysis performed in the present study. As an example, the results for images of regular print (Times New Roman serif font, **Figure 1**; Arno Pro serif font, **Figure 2**) are shown. In the Fourier power spectrum



(Figure 1B), low spatial frequencies are represented at the center and high frequencies at the periphery. Lighter shades represent more spectral power. For each frequency, power was radially averaged in sectors representing cardinal (horizontal and vertical) and oblique orientations (Figure 1C) and plotted as a function of spatial frequency in separate log-log plots (Figures 1D–F).

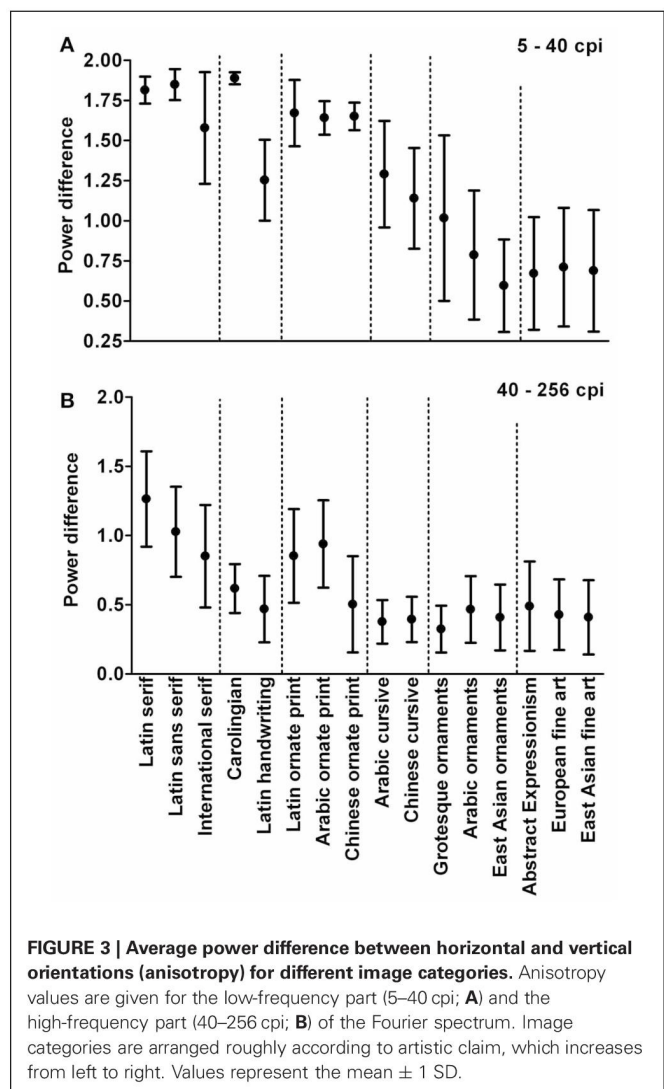
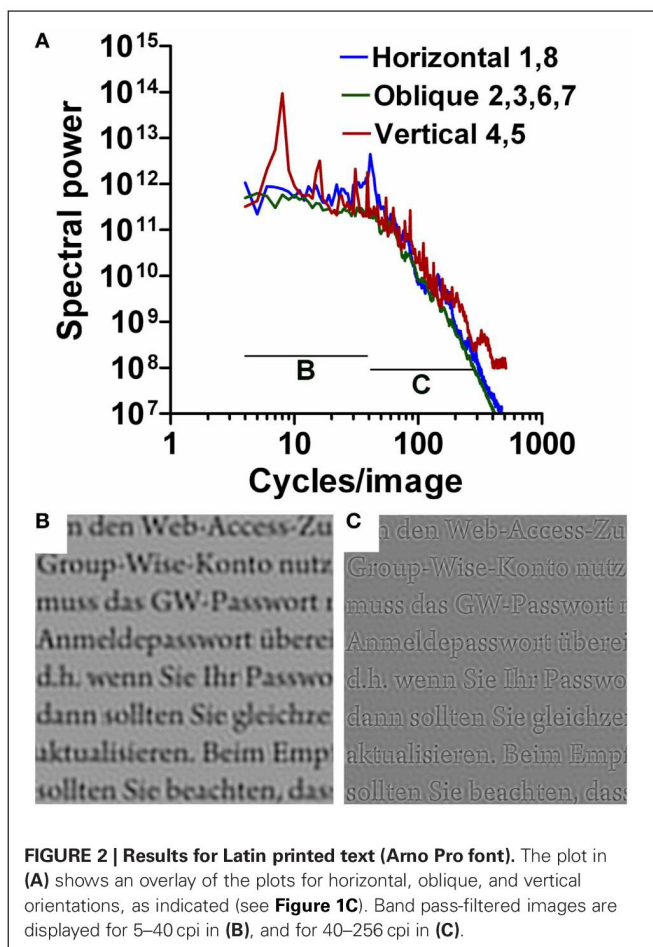
In contrast to similar plots for images of natural scenes or artworks (Burton and Moorhead, 1987; Field, 1987; Tolhurst et al., 1992; Graham and Field, 2007; Redies et al., 2007a,b), the plots for regular text can be roughly divided into two parts (Figure 2A). In the low-frequency part of the spectrum, the average curves for the cardinal and oblique orientations are more shallow; the plot for vertical orientations (red curve in Figure 2A) contains a major peak at about 8 cpi that corresponds to the periodicity of text lines in the images, as expected. To visualize this low-frequency part of the spectrum for the reader, a bandpass-filtered (5–40 cpi) representation is displayed in Figure 2B; it shows a blurred version of the text image. In the high-frequency part, the curves are steeper and fall-off linearly in the log-log plots. In the bandpass-filtered representation of the high-frequency part (40–256 cpi; Figure 2C), the outlines of all letters are sharply demarcated. The two parts of the curves form a transition at around 40 cpi for images with eight lines of text. In control images with four lines and 16 lines of Latin printed text, the transition is shifted to

about 20 and 80 cpi, respectively (data not shown), suggesting that the transition point is found at a spatial frequency about five times higher than the frequency peak that reflects the number of lines.

### ANISOTROPY

Figure 3 shows the mean difference between power values for vertical and horizontal orientations (anisotropy) for all image categories. Results are arranged with subjective artistic claim increasing from left to right, from regular print and handwriting to ornate print, ornamental art, calligraphy, and artworks. As a word of caution, however, we note that a concept like artistic claim is difficult to quantify and may be subject to various philosophical and art historical considerations (see Introduction). Results are presented separately for the low-frequency part (Figure 3A) and the high-frequency part of the spectrum (Figure 3B).

As expected, anisotropy values are high for regular print in both parts of the spectrum. Similarly high values are obtained for ornate print and for Carolingian handwriting that resembles



regular print in its uniform stroke width and regular letter alignment. Values are lower for fine art ( $p < 0.001$ ), confirming previous results (Koch et al., 2010), and for East Asian ornaments (porcelain decorations) and Arabic ornaments (wall decorations) ( $p < 0.001$ ). Significances of the differences between the different image categories were determined by the Tukey range test throughout this work. For the low-frequency part of the spectrum (**Figure 3A**), values for Latin handwriting, calligraphy, and grotesque ornaments are intermediate between artworks and regular print ( $p < 0.001$ ). For the high-frequency part (**Figure 3B**), these image categories are about as lower or higher than those of fine art.

### SLOPE MEASUREMENTS

To quantify the steepness of the curve in the low-frequency part (5–40 cpi) and in the high-frequency part (40–256 cpi), we calculated the slopes of straight lines that were fitted to the curves in the two ranges for each image (for examples, see **Figures 1D–F**). We also measured the differences between the slopes of the high-frequency and low-frequency parts (see Materials and Methods section). For a continuous straight line through both ranges, this difference assumes a value of 0. For each category of images, the two slopes and their difference are listed in **Table 1** for the two cardinal (vertical and horizontal) and the oblique orientations. Moreover, to assess how well these straight lines fitted the curves, the deviation (sigma) of the curves from the fitted lines was also determined. In the following sections, we will describe the results for each of the image categories that were analyzed in the present study.

### General overview

Like in **Figure 3**, results are arranged with artistic claim increasing from left to right in **Figure 4**. In this direction, the slope in the low-frequency part (5–40 cpi; **Figures 4A–C**) becomes more negative and approaches values between around  $-2$  and  $-2.5$ . The slope in the high-frequency part assumes less negative values and approaches similar slope values (40–256 cpi; **Figures 4D–F**). As a result, the difference between the two slopes decreases with increasing artistic claim (**Figures 4G–I**) and approaches values of about 0. These general tendencies are similar for all three orientations ranges.

### Regular print

To study whether different categories of regular printed text result in similar graphs, we carried out the same type of analysis for multiple Latin fonts (serif and sans serif) as well as examples of regular print of other provenances (Arabic, Chinese, and other international fonts). An example of each type of font is displayed in **Figures 5A,D,G**. Averaged curves for all fonts within one category (one example for each font) are shown in **Figures 5B,E,H** and averaged curves for 30 examples of one font are displayed in **Figures 5C,F,I**. The within-font variance was similar or less than the between-font variance (data not shown).

For all plots in **Figure 5**, the steepness of the curve changes at around 40 cpi, similar to the example shown in **Figure 1**. The mean slopes for the low-frequency part ( $-0.78$  to  $0.19$ ) indicate that power is constant or falls less strongly with increasing frequency than for the high-frequency part (mean slopes

between  $-3.71$  and  $-2.91$ ). The slope difference assumes mean values between 2.23 and 3.81 (**Table 1**; **Figure 4**).

### Artworks

To quantify the expected difference between regular print and artworks, we carried out the same type of analysis for two different datasets of artwork images, namely 200 examples of European graphic fine art (Redies et al., 2007b; Graham and Field, 2008), and 209 examples of East Asian monochrome paintings and prints. **Figures 6A, 7A** show examples of the original images analyzed, together with their 2d Fourier power spectra (**Figures 6B, 7B**). With increasing spatial frequency, power falls nearly linearly according to a power law ( $1/f^2$  characteristics) for all orientations (**Figures 6C, 7C**). The band-pass filtered images that correspond to the image in **Figure 6A** are shown in **Figures 6D,E**. Confirming previous studies (Redies et al., 2007b; Graham and Field, 2008), there is no abrupt transition at 40 cpi between the slopes of the low-frequency part and the high-frequency part.

The results listed in **Table 1** confirm that the slopes of art images is more negative ( $p < 0.001$ ; Tukey range test) than images of regular print and assume average values around  $-2$  for the low-frequency part of the spectrum (range  $-1.76$  to  $-2.20$ ). In the high-frequency part, the slope is less negative ( $p < 0.001$ ) with average values between  $-1.86$  and  $-2.67$ , depending on the orientations. The slope difference is around 0 (i.e., close to a straight line; range  $-0.34$  to  $0.86$ ) and differs from printed text ( $p < 0.001$ ; **Table 1**, **Figures 1–5**).

### Handwriting

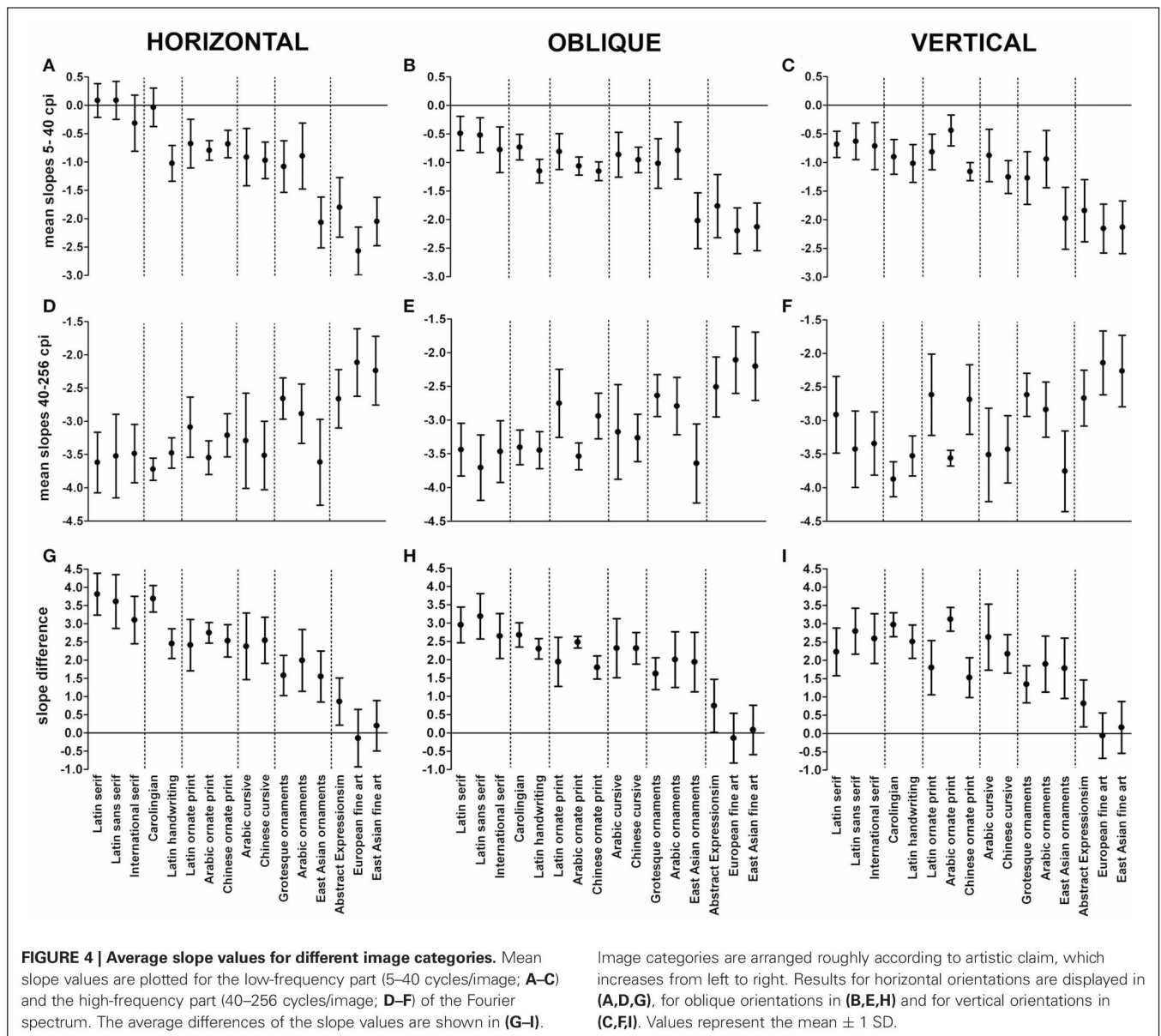
Next, we asked whether samples of handwritten text share some of the features observed for regular print. We analyzed two types of historical Latin handwriting. **Figure 8** show results from a Carolingian manuscript (10th century; **Figures 8A–C**) and for handwritten Latin text from the Reformation period (dated 1528; **Figures 8D–F**), respectively. For Carolingian writing, overall results are similar to those of regular print. Similar to regular print, Carolingian handwriting is characterized by a uniform stroke width and a regular vertical alignment and spacing of the letters. For Latin handwriting, slopes are more negative in the low-frequency part (range  $-1.02$  to  $-1.15$ ;  $p < 0.001$ ) and the slope difference is smaller for horizontal and oblique orientations than for regular print (2.45 and 2.30, respectively;  $p < 0.001$ ; **Table 1**; **Figure 4**).

### Ornate print and calligraphy

We also analyzed writing systems with artistic claim. Firstly, we analyzed ornate fonts of three different cultural provenances (Latin, Arabic, and Chinese). Results for Latin and Arabic ornate print are shown in **Figure 9**. Secondly, we analyzed Arabic and Chinese cursive calligraphy. Typically, the cursive Arabic letters have long ascenders and descenders, while cursive Chinese pictograms are characterized by quick, fluent brushwork. **Figure 10** shows results for Chinese calligraphy.

Similar to the samples of Latin handwriting, slope values for the horizontal direction (range  $-0.67$  to  $-0.99$ ) and vertical direction (range  $-0.44$  to  $-1.25$ ) are smaller than for regular



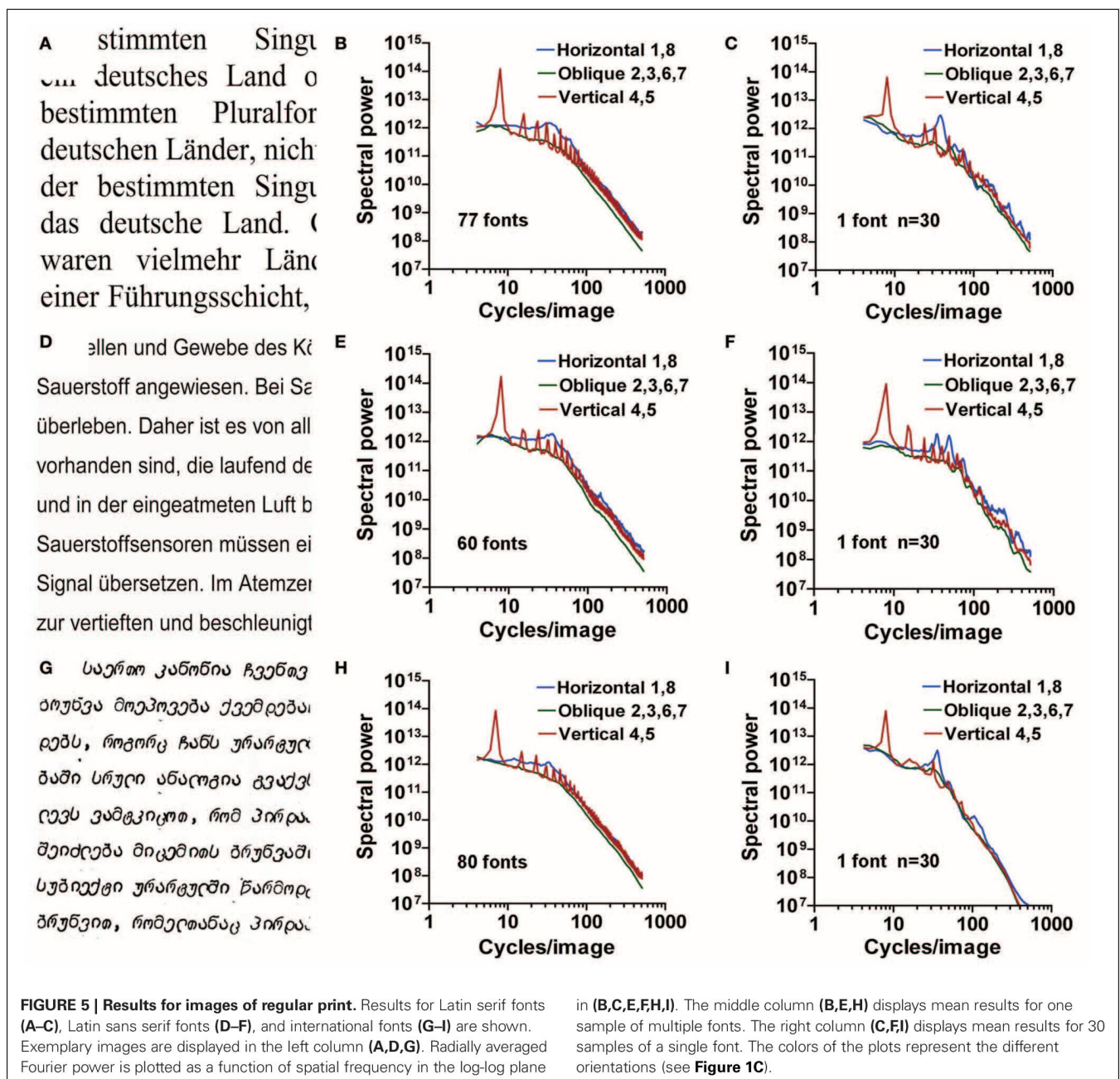


print ( $p < 0.001$ ) and higher than those for aesthetic artworks ( $p < 0.001$ ) in the low-frequency part of the spectrum. An exception is Arabic ornate print with values that are similar to regular print for the slopes of the high-frequency part and the slope difference. Results for the other orientations and parts of the spectrum are similar to those of regular print.

### Ornamental art

The arrangement of letters in regular text without artistic claim can be described as a sequence of largely independent pictorial elements. In contrast, artworks are characterized by a more global composition, in which individual pictorial elements relate to each other throughout the entire image (see Introduction). We next studied images of ornamental art (Western, Arabic, and East Asian), which consist also of repetitive pictorial elements arranged in a global structure (Figures 11–13).

The mean log-log plots for the three datasets are characterized by a more gradual rather than an abrupt transition of Fourier power at around 40 cpi (Figures 11C, 12C, 13C). Results for Western grotesque ornaments (Figure 11) and ornamental Arabic art (wall decorations; Figure 12) are similar to those of ornate print and calligraphy in general. For the low-frequency part of the spectrum, the slope values for ornamental paintings on East Asian porcelain (Figure 13) are lower than for Western and Arabic ornamentals ( $p < 0.001$ ) and resemble those of aesthetic artworks (Figures 4A–C). The opposite tendency is observed for the high-frequency part of the spectrum (Figures 4D–F;  $p < 0.001$ ). Here, slope values for East Asian decorative art are similar to those of regular print and lower than those of Western and Arabic decorative art ( $p < 0.001$ ). For all orientation ranges, slope differences (Figures 4G–I) are equal or higher ( $p < 0.001$ ) for ornamental art than for images of ornamental writing or calligraphy.



### Abstract expressionism

Finally, particular styles of abstract art can be described as an arrangement of similar pictorial elements embedded in a global image structure, similar to ornamental art. In the present work, we analyzed paintings by Abstract Expressionist artists (Jackson Pollock, Jean Dubuffet, Cy Twombly, and Christian Dotremont). Examples of the images cannot be shown for copyright reasons.

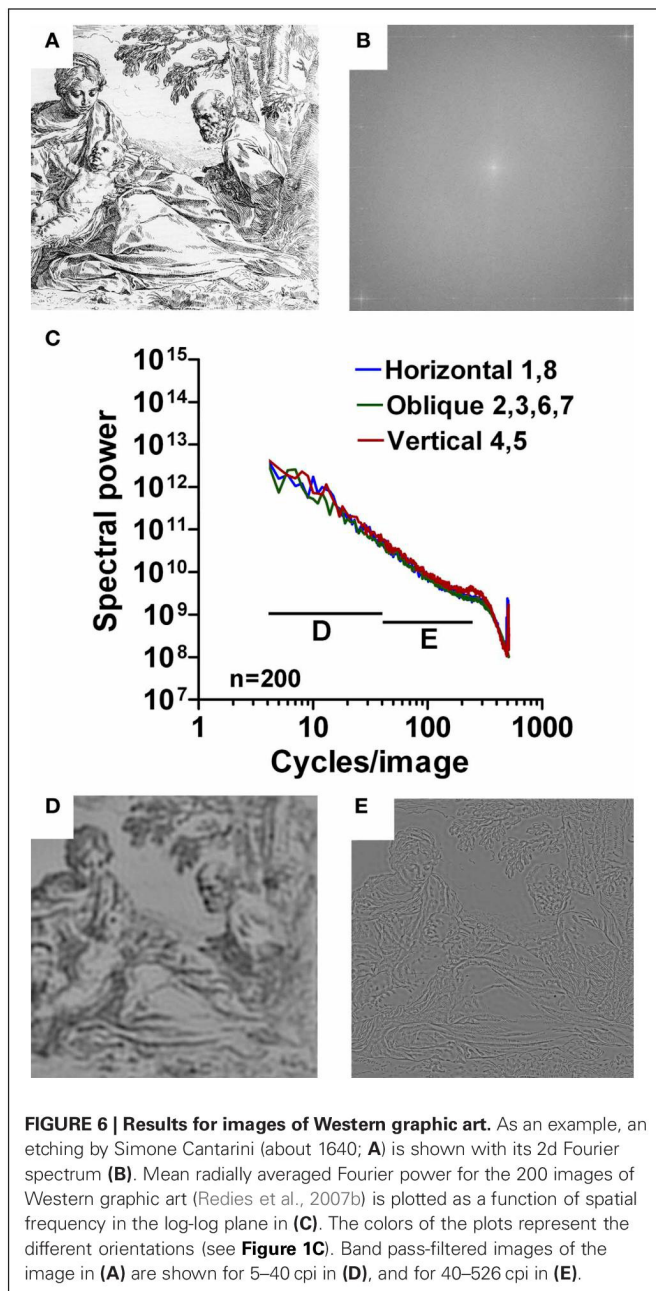
The mean log-log plot for the Abstract Expressionist dataset (not shown) is similar to the curves for fine art. The slope values in the low-frequency range are around  $-1.8$  (Table 1; Figures 4A–C), which is only slightly lower than the value for European and Asian fine art and East Asian porcelain decorations

(around  $-2.1$ ,  $p < 0.001$ ), but lower ( $p < 0.001$ ) than the mean value for artistic Western and Arabic ornaments (around  $-1.0$ ). In the high-frequency part (Figures 4D–F), however, the curve for abstract expressionism assumes slope values of around  $-2.6$ . This is lower ( $p < 0.001$ ) than the value for fine art ( $-1.9$ ). The slope differences (Figures 4G–I) are around  $0.74$ – $0.86$ , compared to  $1.34$ – $2.04$  for artistic ornaments ( $p < 0.001$ ) and  $-0.34$ – $-0.19$  for fine art ( $p < 0.001$ ).

### DISCUSSION

In the present work, we compared image statistics of ordinary text and different categories of images with artistic claim. Humans

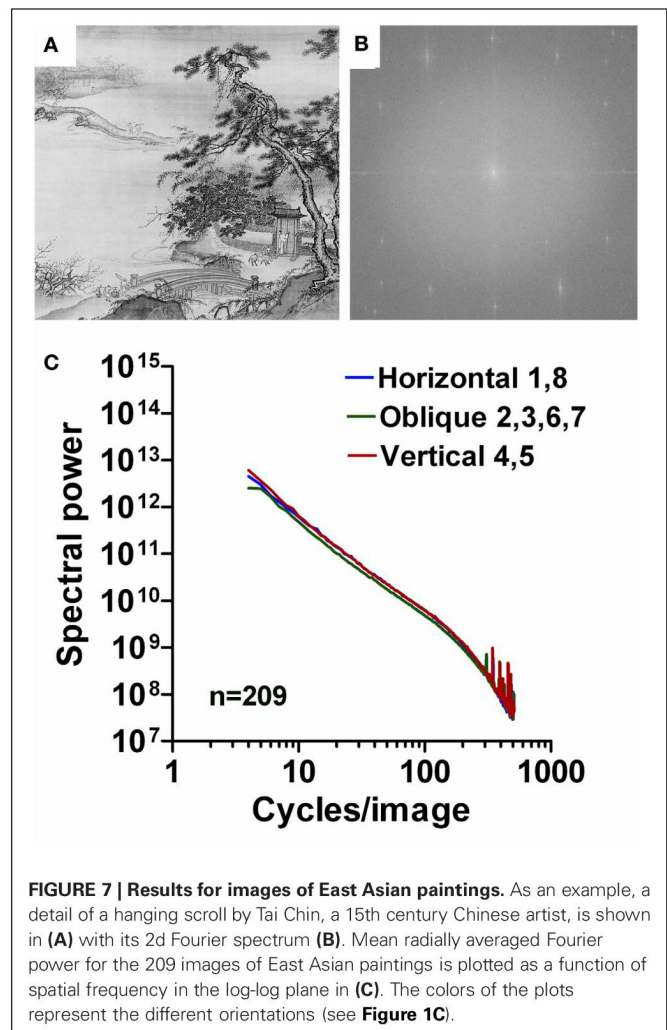




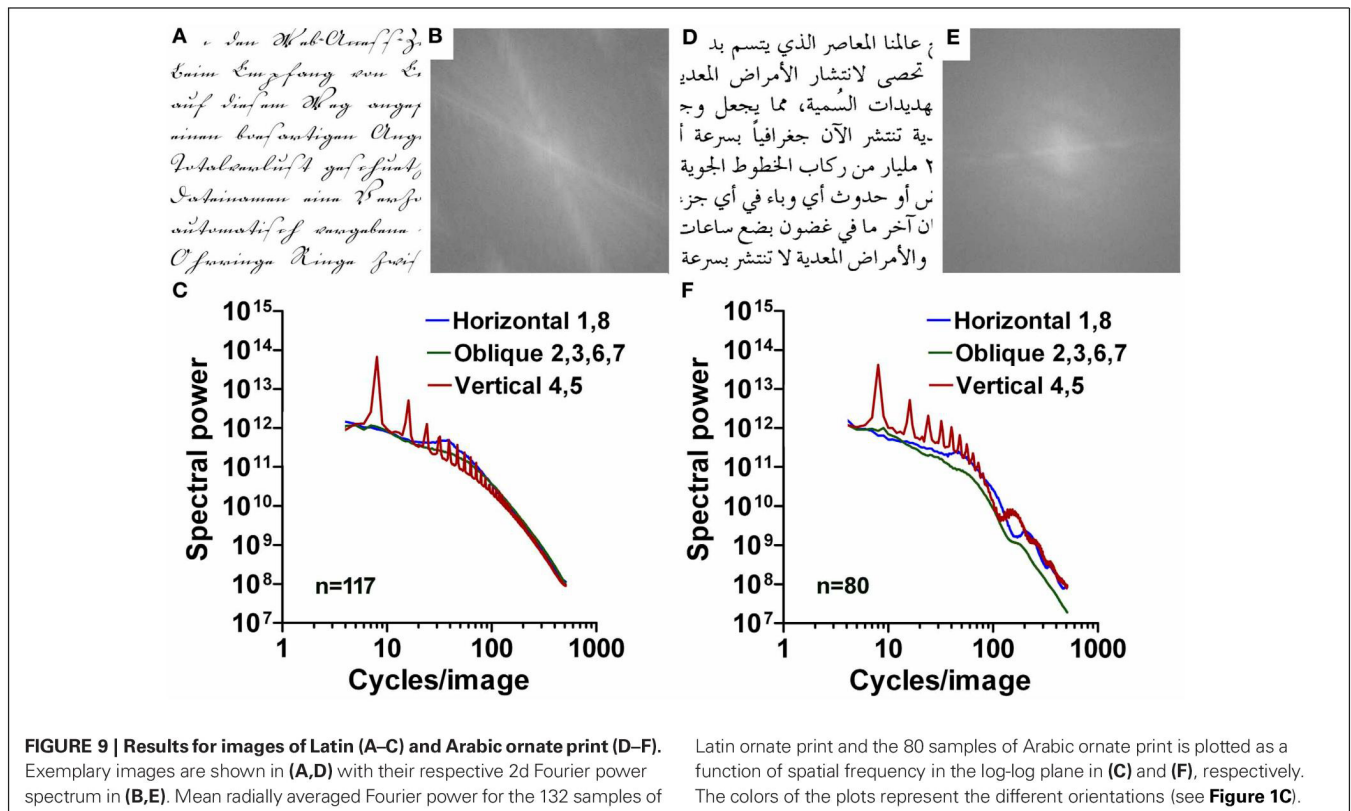
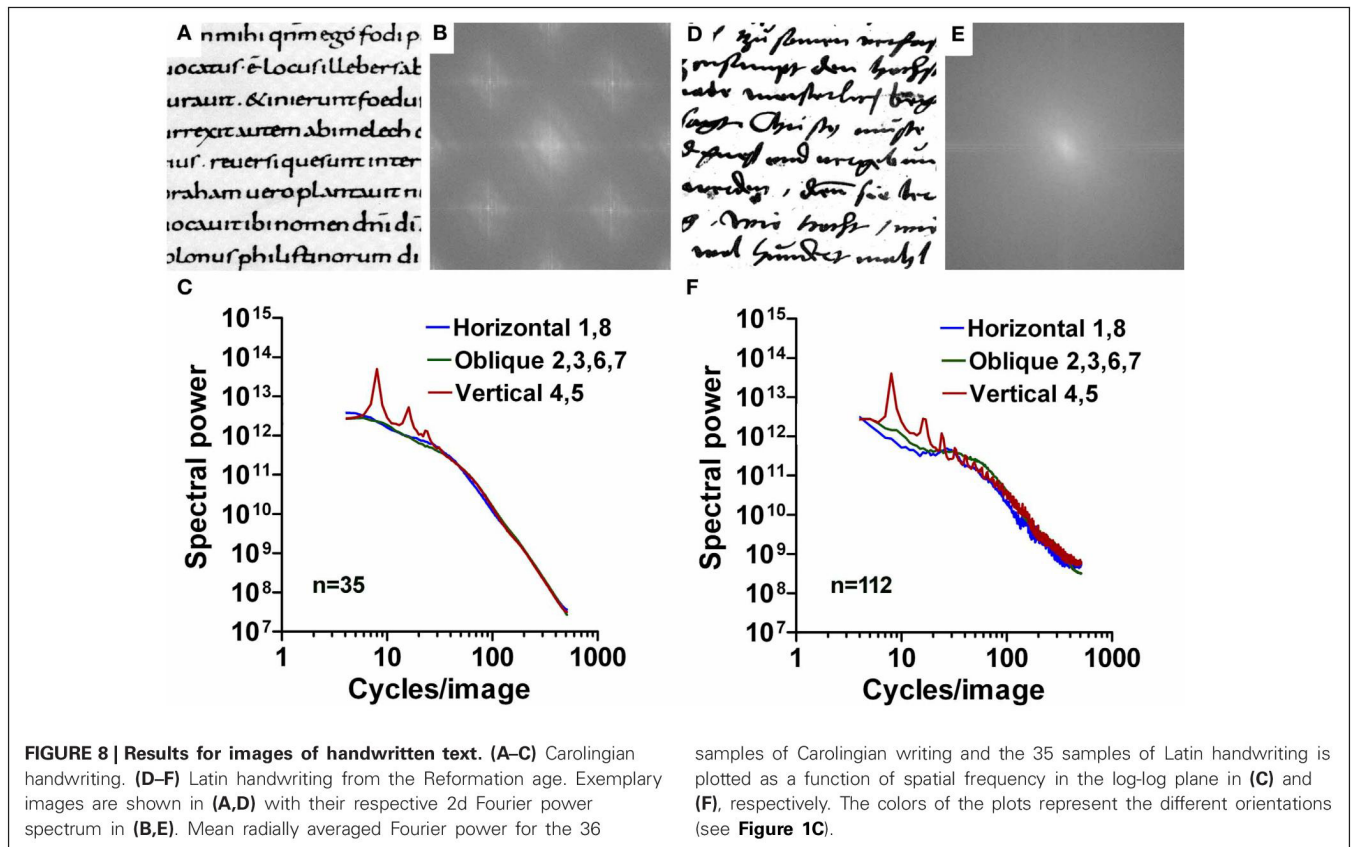
create all these images for viewing by humans. As a consequence, the images may exhibit statistical properties that reflect sensory adaption to the human visual system (Changizi et al., 2006; Graham and Redies, 2010). However, the aesthetic appeal and artistic intent of the image categories differ (see Introduction). Our results reveal that, in general, specific statistical properties vary with the artistic claim of the images.

#### IMAGES OF REGULAR PRINT ARE NOT SCALE-INVARIANT

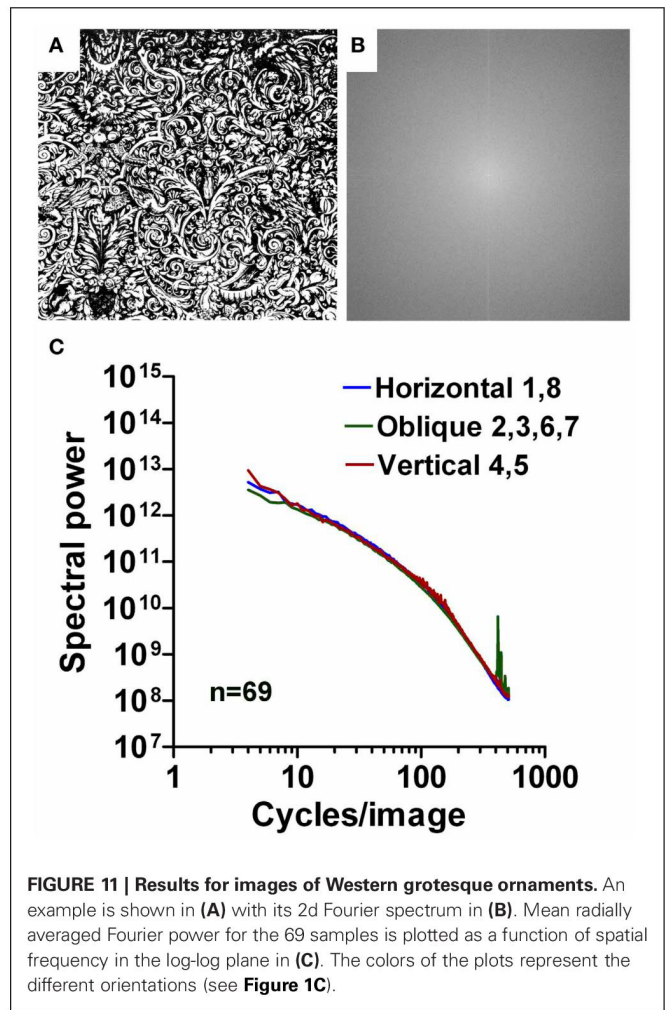
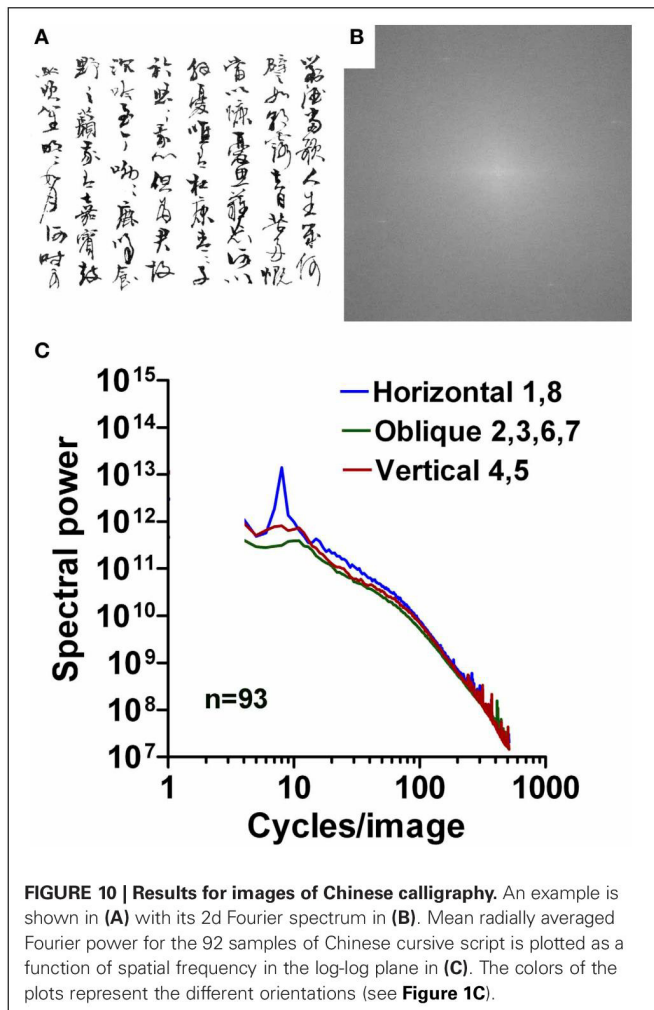
It can be expected that aesthetic artworks and regular text differ in their Fourier power spectra (see Introduction). In the present work, we provide a systematic study of this assumption and



quantify the differences by applying a computer-based algorithm for measuring statistical image properties. For aesthetic artworks, radially averaged spectral power falls off roughly linearly according to a power law ( $1/f^p$  characteristics) with increasing spatial frequency in log-log plots; the mean slope value [ $p$ ] is about  $-2$  (Graham and Field, 2007; Redies et al., 2007a,b; Graham and Redies, 2010). This result implies that the Fourier spectrum is scale-invariant. Artworks share this property with other types of aesthetically pleasing images (for example, graphic novels and comics; Koch et al., 2010) and with images of complex natural scenes (see Introduction). Fractal-like structure was also found in particular types of music (Voss and Clarke, 1975; Beauvois, 2007), architecture (Joye, 2007), and American sign language (Bosworth et al., 2006). Unlike artworks, regular print has a steeper slope in the high-frequency part of the power spectrum (value of about  $-3.5$ ) while the slope of the low-frequency part is shallower (value of about  $-1.2$ , **Table 1**, **Figure 4**). The difference in the slope values between the three orientation ranges (horizontal, oblique, and vertical) is small. Similar findings for regular handwritten text suggest that, in images of regular text, power in the low-frequency part of the spectrum is relatively low





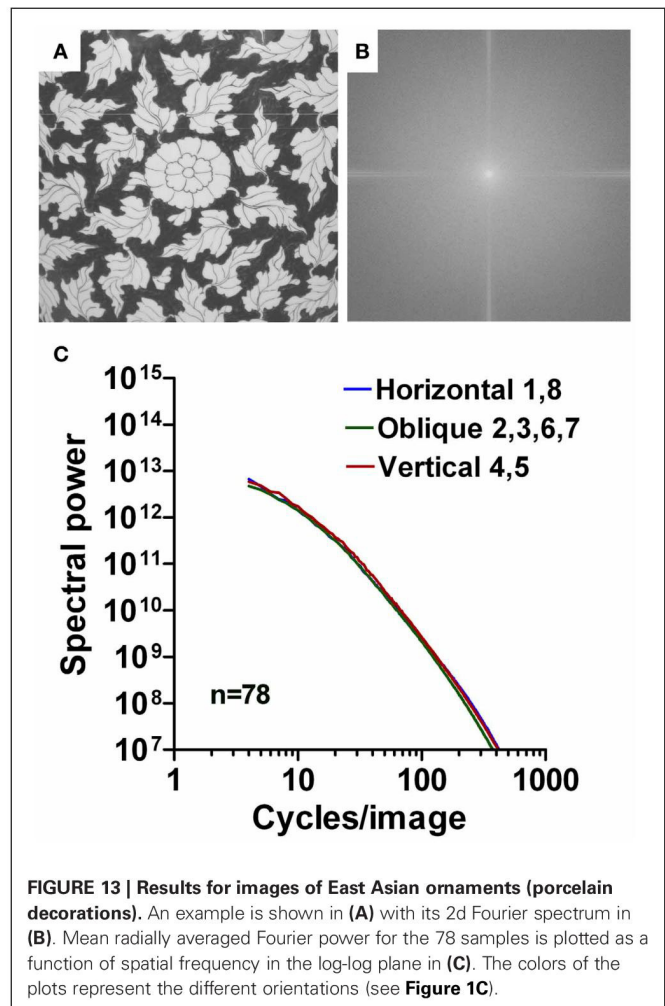
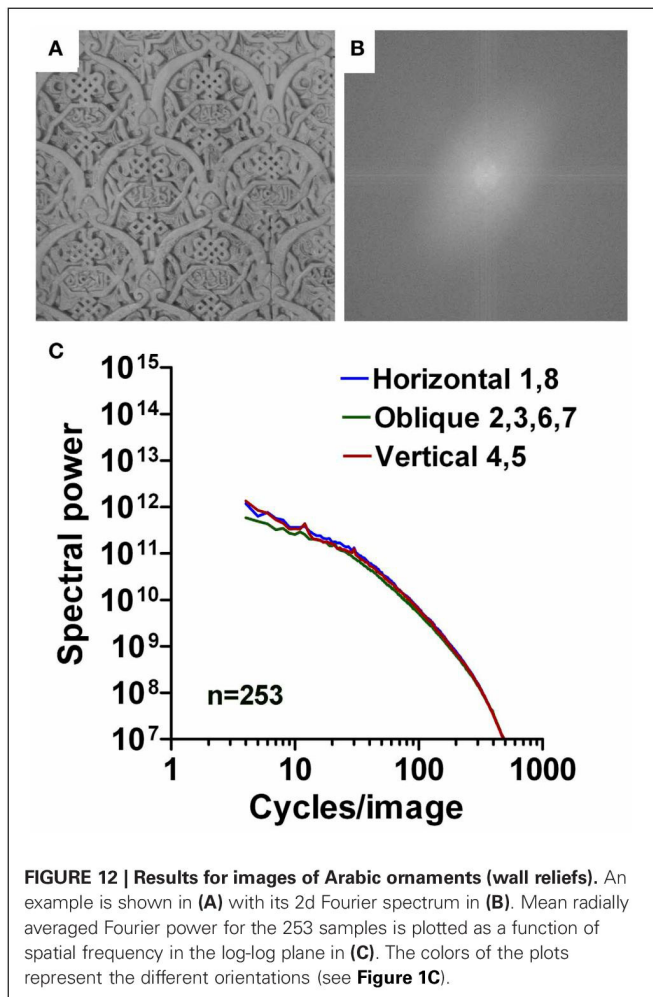


when compared to artworks, with the exception of a prominent peak at 8 cpi (see above). Because low spatial frequencies represent coarse structure in an image and high spatial frequencies represent fine detail, this result implies that images of ordinary text tend to contain a lower amount of global image structure than the artworks analyzed. In the artworks, the higher amount of global structure may represent a physical correlate of artistic composition, which relates individual pictorial elements to each other across the image (see Introduction).

It has been argued that all images produced by the human hand, including artworks, generally possess scale-invariant properties for reasons related to the nature of hand movements (e.g., see Graham and Field, 2008). The present findings indicate that not all hand-made images are scale-invariant in the Fourier domain. The result that images of handwriting contain a lower amount of global image structure is not surprising because global image structure is not important for text, which is scanned word-by-word in a consecutive manner. In addition, even the high-frequency part of the Fourier power spectra of text is not scale-invariant. We conclude that humans can manufacture images that may or may not be scale-invariant, depending on their purpose.

Studies on artificial images revealed that manipulations of the Fourier spectrum can elicit visual discomfort if a significant deviation from scale invariance is induced (Fernandez and Wilkins, 2008; Juricevic et al., 2010; O'Hare and Hibbard, 2011). In particular, images with a curvilinear amplitude spectrum and an energy upshot at about 3–4 cycles per degree, i.e., close to the position where the visual system is most sensitive, can elicit visual discomfort (Fernandez and Wilkins, 2008). Visual discomfort is often (but not always) negatively associated with artistic merit (Fernandez and Wilkins, 2008). In the present study, curvilinear power spectra are also observed for several of the image categories (for example, images of text, calligraphy, and ornamental art), but the images used by us do not evoke obvious visual discomfort in general. Consequently, it remains unclear how the findings by Fernandez and Wilkins (2008) relate to our present results.

Together, these results suggest that several types of images, which lack scale invariance, are of relatively low aesthetic value. The opposite notion, however, does not hold because images, which possess scale-invariant Fourier spectra, are not necessarily aesthetic (for example, some computer-generated images; Lee and Mumford, 1999). It has therefore been suggested that scale invariance is a corollary of some other (unknown) feature of



aesthetic images that contributes to aesthetic perception (Redies et al., 2007a).

#### SLOPE DIFFERENCE AND ANISOTROPY CORRELATE WITH ARTISTIC CLAIM

Images of text are of particular interest for studying aesthetic perception because there is a continuous transition from regular print to aesthetic writing (ornamental writing and calligraphy) and to visual art (see Introduction). Our results (Table 1, Figures 3, 4) illustrate that, with increasing artistic claim, images of text acquire specific statistical properties that are similar to those of visual art. As one such measure, we introduced the difference of the slopes between the low- and high-frequency parts of the radially averaged power spectrum. With increasing artistic claim, this difference decreases to values close to 0 (straight line) for all orientation ranges (Table 1, Figure 4), which indicates a nearly linear fall-off throughout the entire frequency spectrum, similar to what has been observed for natural scenes ( $1/f^2$  characteristics). A similar transition between non-aesthetic images and images with artistic claim is observed for the horizontal/vertical power difference (anisotropy measure; Figure 3). For images of text, differences between vertical and horizontal

orientations probably reflect the periodicity of the text lines and/or regularities in the width or spacing of the lines that form the letters. Compared to regular text, the Fourier spectrum of fine art represents all orientations at similar strength, as shown previously by Koch et al. (2010) who compared artworks to other categories of images. This result is not trivial because artists could easily produce paintings, in which particular orientations predominate. In how far low anisotropy is necessary or sufficient to induce aesthetic perception—and in which types of fine art—remains to be studied.

In our analysis, we included a special style of art (Abstract Expressionism) that resembles text images in its repetitive arrangement of multiple and simple pictorial elements distributed over a large surface area. Nevertheless, the statistical properties of Abstract Expressionist paintings are more similar to those of other art images than to ordinary text, suggesting that they contain a high amount of global structure, similar to fine art. This result is also not trivial because, conceivably, paintings using the same pictorial elements could be produced with a lower amount of global structure. Our results are compatible with the suggestion by R. Taylor and other researchers that abstract expressionist paintings, like the drip paintings by Jackson



Pollock, possess fractal-like structure (Taylor, 2002; Mureika, 2005; Alvarez-Ramirez et al., 2008).

Other examples of repetitive structures arranged over a large surface are ornamental decorations, which are also created to be enjoyed by human observers but may perhaps have lesser artistic claim in general. The slope differences for all three types of decorative art are intermediate between those of fine art and ordinary text. The anisotropy values of decorative art are intermediate or closer to fine art, when compared to regular print (Figure 3). Intermediate values are also obtained for calligraphy, a writing style with artistic claim. It remains unclear whether other types of images show a similar relation between statistical regularities and artistic claim.

### SIMILAR FINDINGS IN TEXT IMAGES OF DIFFERENT CULTURAL PROVENANCE

To compare Western, Arabic, and East Asian examples of the different categories of text images (Table 1, Figures 3, 4), we chose a horizontal approach and compared contemporary text images that include plain and ornate fonts as well as serif and sans serif fonts of different international alphabets. In addition, we chose a vertical approach and analyzed text of different ages (medieval manuscripts, calligraphy, and Reformation handwriting). Overall, we analyzed 11 datasets, each including between 13 and 253 samples (1598 images in total; Table 1).

Results from the Fourier analysis were generally similar for the three cultural backgrounds (Table 1, Figures 3, 4). In particular, the slope differences were similar for Latin and international serif fonts of regular print, and also for ornate print, calligraphy, and ornamental art from all three cultures. Moreover,

our results confirm that artworks from both East Asian and Western provenance possess similar scale-invariant properties in the Fourier domain (Redies et al., 2007b; Graham and Field, 2008). Together, these results suggest that specific perceptual mechanisms for reading and aesthetic judgment, respectively, may be common amongst humans across different cultural backgrounds. It has been speculated that such common principles may have emerged due to selective pressures imposed by the adaptation of the human visual system to specific perceptual and motor tasks during the evolution of mankind (Changizi and Shimojo, 2005; Changizi et al., 2006; Redies, 2007; Graham and Redies, 2010).

Last but not least, physical features of the visual inputs have been shown to strongly modulate the functional responses in some core regions of the reading network, including for instance the influence of spatial frequency on the activation of the left ventral occipitotemporal cortex (Seghier and Price, 2011; Woodhead et al., 2011; Horie et al., 2012). The kind of image statistics studied here may thus shed some light on how the human brain processes written word stimuli in comparison to other types of stimuli. This topic warrants future studies.

### ACKNOWLEDGMENTS

The authors thank Mrs. Lisa Redies for producing and scanning of printed text, Mrs. Julia Braun for assistance with the statistical analysis, and members of the Denzler and Redies groups for constructive suggestions, discussion, and comments on the manuscript. They are grateful to Prof. Dr. Ulrich Pietsch for permission to take photographs at the Dresden Porcelain Collection, Dresden, Germany.

### REFERENCES

- Alvarez-Ramirez, J., Ibarra-Valdez, C., Rodriguez, E., and Dagdug, L. (2008). 1/f-noise structures in Pollock's drip paintings. *Phys. A* 387, 281–295.
- Beauvois, M. W. (2007). Quantifying aesthetic preference and perceived complexity for fractal melodies. *Music Percept.* 24, 247–264.
- Bosworth, R. G., Bartlett, M. S., and Dobkins, K. R. (2006). Image statistics of American Sign Language: comparison with faces and natural scenes. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 23, 2085–2096.
- Burton, G. J., and Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Appl. Phys.* 26, 157–170.
- Changizi, M. A., and Shimojo, S. (2005). Character complexity and redundancy in writing systems over human history. *Proc. Biol. Sci.* 272, 267–275.
- Changizi, M. A., Zhang, Q., Ye, H., and Shimojo, S. (2006). The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *Am. Nat.* 167, E117–E139.
- Chung, S. T., and Tjan, B. S. (2007). Shift in spatial scale in identifying crowded letters. *Vision Res.* 47, 437–451.
- Chung, S. T., and Tjan, B. S. (2009). Spatial-frequency and contrast properties of reading in central and peripheral vision. *J. Vis.* 9, 16.1–16.19.
- Fernandez, D., and Wilkins, A. J. (2008). Uncomfortable images in art and nature. *Perception* 37, 1098–1113.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 4, 2379–2394.
- Graham, D., and Redies, C. (2010). Statistical regularities in art: relations with visual coding and perception. *Vision Res.* 50, 1503–1509.
- Graham, D. J., and Field, D. J. (2007). Statistical regularities of art images and natural scenes: spectra, sparseness and nonlinearities. *Spat. Vis.* 21, 149–164.
- Graham, D. J., and Field, D. J. (2008). Variations in intensity statistics for representational and abstract art, and for art from the Eastern and Western hemispheres. *Perception* 37, 1341–1352.
- Horie, S., Yamasaki, T., Okamoto, T., Kan, S., Ogata, K., Miyauchi, S., et al. (2012). Distinct role of spatial frequency in dissociative reading of ideograms and phonograms: an fMRI study. *Neuroimage* 63, 979–988.
- Hoyer, P. O., and Hyvärinen, A. (2002). Sparse coding of natural contours. *Neurocomputing* 44–46, 459–466.
- Jainta, S., Jaschinski, W., and Wilkins, A. J. (2010). Periodic letter strokes within a word affect fixation disparity during reading. *J. Vis.* 10, 2.
- Joye, Y. (2007). Architectural lessons from environmental psychology: the case of biophilic architecture. *Rev. Gen. Psychol.* 11, 305–328.
- Juricevic, I., Land, L., Wilkins, A., and Webster, M. A. (2010). Visual discomfort and natural image statistics. *Perception* 39, 884–899.
- Klemm, E. (1998). *Die Illuminierten Handschriften des 13. Jahrhunderts*: Teil 4. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Koch, M., Denzler, J., and Redies, C. (2010). 1/f 2 Characteristics and isotropy in the fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs. *PLoS ONE* 5:e12268. doi: 10.1371/journal.pone.0012268
- Lee, A. B., and Mumford, D. (1999). “An occlusion model generating scale-invariant images,” in *Proceedings of the IEEE Workshop on Statistical and Computer Theories of Vision* (Fort Collins, CO), 1–20.
- Majaj, N. J., Pelli, D. G., Kurshan, P., and Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Res.* 42, 1165–1184.
- Menz, G. (1912). *Handschriften der Reformationszeit*. Bonn: Markus und Weber.
- Mureika, J. R. (2005). Fractal dimensions in perceptual color space: a comparison study using Jackson Pollock's art. *Chaos* 15:043702. doi: 10.1063/1.2121947
- Nandy, A. S., and Tjan, B. S. (2008). Efficient integration across spatial

- frequencies for letter identification in foveal and peripheral vision. *J. Vis.* 8, 3.1–3.20.
- O'Hare, L., and Hibbard, P. B. (2011). Spatial frequency and visual discomfort. *Vision Res.* 51, 1767–1777.
- Olshausen, B. A., and Field, D. J. (1996). Natural image statistics and efficient coding. *Network* 7, 333–339.
- Oruc, I., and Landy, M. S. (2009). Scale dependence and channel switching in letter identification. *J. Vis.* 9, 4.1–4.19.
- Ozawa, S. (1994). “1/f fluctuation in Japanese brush calligraphy,” in *Research of Pattern Formation*, ed R. Takaki (Tokyo: KTK Scientific Publishers), 571–579.
- Parraga, C. A., Troscianko, T., and Tolhurst, D. J. (2000). The human visual system is optimised for processing the spatial information in natural visual images. *Curr. Biol.* 10, 35–38.
- Pelli, D. G., Burns, C. W., Farell, B., and Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Res.* 46, 4646–4674.
- Quiroga, R. Q., and Pedreira, C. (2011). How do we see art: an eye-tracker study. *Front. Hum. Neurosci.* 5:98. doi: 10.3389/fnhum.2011.00098
- Redies, C. (2007). A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spat. Vis.* 21, 97–117.
- Redies, C., Hänisch, J., Blickhan, M., and Denzler, J. (2007a). Artists portray human faces with the Fourier statistics of complex natural scenes. *Network* 18, 235–248.
- Redies, C., Hasenstein, J., and Denzler, J. (2007b). Fractal-like image statistics in visual art: similarity to natural scenes. *Spat. Vis.* 21, 137–148.
- Seghier, M. L., and Price, C. J. (2011). Explaining left lateralization for words in the ventral occipitotemporal cortex. *J. Neurosci.* 31, 14745–14753.
- Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Curr. Opin. Neurobiol.* 13, 144–149.
- Solomon, J. A., and Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature* 369, 395–397.
- Taylor, R. P. (2002). Order in Pollack's chaos – Computer analysis is helping to explain the appeal of Jackson Pollock's paintings. *Sci. Am.* 287, 116–121.
- Tolhurst, D. J., Tadmor, Y., and Chao, T. (1992). Amplitude spectra of natural images. *Ophthal. Physiol. Opt.* 12, 229–232.
- Tyler, C. W., and Likova, L. T. (2007). Crowding: a neuroanalytic approach. *J. Vis.* 7, 16.1–16.9.
- Vinje, W. E., and Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287, 1273–1276.
- Voss, R. F., and Clarke, J. (1975). 1/f noise in music and speech. *Nature* 258, 317–318.
- Warncke, C.-P. (1979). *Die Ornamentale Groteske in Deutschland 1500-1650: Volume 2*. Berlin: Verlag Volker Spiess.
- Wilkins, A. J. (1995). *Visual Stress*. Oxford: Oxford University Press.
- Wilkins, A. J., Smith, J., Willison, C. K., Beare, T., Boyd, A., Hardy, G., et al. (2007). Stripes within words affect reading. *Perception* 36, 1788–1803.
- Woodhead, Z. V., Wise, R. J., Sereno, M., and Leech, R. (2011). Dissociation of sensitivity to spatial frequency in word and face preferential areas of the fusiform gyrus. *Cereb. Cortex* 21, 2307–2312.
- Wooding, D. S., Mugglestone, M. D., Purdy, K. J., and Gale, A. G. (2002). Eye movements of large populations: I. Implementation and performance of an autonomous public eye tracker. *Behav. Res. Methods Instrum. Comput.* 34, 509–517.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 February 2013; accepted: 13 March 2013; published online: 01 April 2013.

Citation: Melmer T, Amirshahi SA, Koch M, Denzler J and Redies C (2013) From regular text to artistic writing and artworks: Fourier statistics of images with low and high aesthetic appeal. *Front. Hum. Neurosci.* 7:106. doi: 10.3389/fnhum.2013.00106

Copyright © 2013 Melmer, Amirshahi, Koch, Denzler and Redies. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

## DISKUSSION

Bereits seit den 1930er Jahren werden Kunstwerke in der Sehforschung verwendet, um die menschliche Wahrnehmung zu erforschen (z.B. Buswell, 1935). Diese methodische Strategie erweist sich als zielführend, weil in Kunstbildern verschiedene optische Verfahren zum Einsatz kommen, um beispielsweise Räumlichkeit oder die Beschaffenheit verschiedener Oberflächen zur Darstellung zu bringen. Zum Beispiel kodieren Künstler bei der Schaffung ihrer Werke die wesentlichen Tiefenkriterien ihrer dreidimensionalen Lebenswelt um, indem sie diese auf eine zweidimensionale Fläche übertragen. Intuitiv nutzen sie dabei Verfahren, die es dem Betrachter erlauben, die tatsächlichen räumlichen Gegenebenenheiten auf Basis dieser flächigen Darstellung zu rekonstruieren. Anhand dieser Überlegung wird deutlich, dass Kunstbilder dem menschlichen Sehsystem gezielt Stimuli anbieten, die von diesem besonders effizient verarbeitet werden können. Im Rahmen von *Eye-Tracking*-Studien wurde bereits der spezifische Pfad der Augenbewegungen untersucht, der von Betrachtern gewählt wird, während diese die Informationen der Bildfläche wahrnehmen (Wooding et al 2001; Quiroga und Pedreira, 2001). Grundvoraussetzung für die Hinwendung der Aufmerksamkeit des Betrachters ist die in der Einleitung erwähnte *inherent interestingness* (Hochberg, 1978), die Darstellungen mit künstlerischem Anspruch gemein zu sein scheint.

Der Aspekt der inhärenten Interessantheit geht über rein neurophysiologische Fragestellungen hinaus und rückt zusätzlich wahrnehmungspsychologische Aspekte in den Fokus. Von besonderem Interesse ist dabei die Frage nach jenen Bildinformationen, die nicht nur die Aufmerksamkeit der Betrachter erregen, sondern sogar eine ästhetische Erfahrung in ihnen hervorrufen können. Die Erforschung der statistischen Eigenschaften von Kunstbildern, wie sie erst seit wenigen Jahren praktiziert wird (siehe Einleitung), erweist sich in diesem Zusammenhang als richtungsweisend. Während bei der Erforschung dieser Eigenschaften einfache Merkmale wie Farbigkeit oder Helligkeit in Betracht gezogen werden können, untersuchte meine Arbeit vorrangig die globale räumliche Struktur der untersuchten Bildkategorien.

In der vorliegenden Arbeit wurden auch die statistischen Eigenschaften verschiedener Textkategorien mit Bildern verglichen, die einen gewissen künstlerischen Anspruch erheben. Allen Bildern ist gemein, dass sie von Menschen geschaffen wurden, um von Menschen betrachtet zu werden. Aus diesem Grund ist es wahrscheinlich, dass diese Bilder statistische

Eigenschaften besitzen, die der Funktionsweise bzw. den Adaptationsmechanismen des menschlichen Sehsystems in besonderem Maße gerecht werden (Changizi et al 2006; Graham und Redies, 2010). Gleichzeitig erheben die untersuchten Bildkategorien unterschiedliche künstlerische und ästhetische Ansprüche, wie in der Einleitung erwähnt wurde.

Die Ergebnisse meiner Arbeit zeigen, dass der ästhetische Anspruch einer Bildkategorie mit ihren spezifischen statistischen Eigenschaften korreliert. Diese Korrelation beweist meiner Meinung nach die zentrale Funktion der ästhetischen Verfremdung als notwendige Voraussetzung der ästhetischen Erfahrung. Das folgende Kapitel soll den Bezug zwischen der Wahrnehmung ‚inhärent interessanter Zeichen‘ und der daraus resultierenden Ästhetik sowohl aus geisteswissenschaftlicher als auch aus naturwissenschaftlicher Sicht nachvollziehen.

## 1. Kunst und ästhetische Erfahrung: Interdisziplinäre Definitionen und Voraussetzungen

I shall argue that art – or, to use a more exact phrase, aesthetic experience – is an essential factor in human development, and, indeed a factor on which Homo sapiens has depended for the development of his highest cognitive faculties.

(Read, 1954)

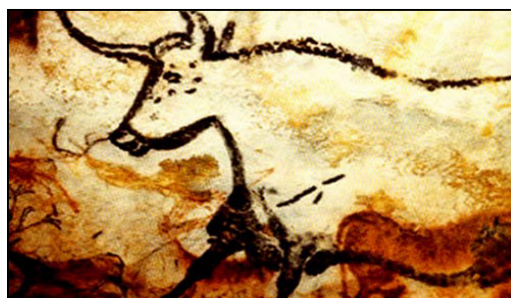


Abb. 4 - Ausschnitt der Höhlenmalereien von Lascaux (Wikimedia Commons, <http://fr.wikipedia.org/wiki/Image:Lascaux-salle-des-taureaux.jpg>, 05.08.2013)

Für den britischen Dichter, Philosophen und Kunstkritiker Sir Herbert Edward Read sind die Höhlenmalereien von Lascaux (vgl. Abb.4) nicht nur das erste Kunstwerk der Menschheitsgeschichte, sondern auch der früheste Beleg einer ästhetischen Erfahrung. Wie

Read in seinem Essay *Art and the Evolution of Consciousness* (1954) ausführt, handelt es sich bei den stilistischen Merkmalen der berühmten Felsenzeichnung keineswegs um eine primitive Frühform künstlerischen Ausdrucks, die im Zuge der fortschreitenden kognitiven Entwicklung des Menschen überwunden wurde und inzwischen als ‚überholt‘ gelten muss. Vielmehr weist er drauf hin, dass die Jagdszenen von Lascaux eine Reihe stilistischer Merkmale enthalten, die sich beispielsweise bei Picasso wiederfinden, der laut Augenzeugen seinerseits bei der Besichtigung der Höhle von Lascaux feststellte: „We have invented nothing“. Für Read zeigen sich in den Jagdszenen *konstante Faktoren* (Read, 1954), die sich durch sämtliche Epochen der Kunstgeschichte ziehen und die offenbar seit jeher mit künstlerischen Ausdrucksformen verknüpft sind.

Dieser Auffassung von Kunst waren bereits in den 1920er Jahren die Vertreter der philosophischen Schule des Prager Strukturalisten um Roman Jakobson (1992). Beeinflusst von Ferdinand de Saussures Zeichenlehre gingen die Strukturalisten davon aus, dass Kunst über eine spezifische Komposition verfügt, die sich formal von der Struktur alltäglicher Stimuli unterscheidet und dadurch ästhetische Erfahrungen auslöst. Diese als *Verfremdungseffekt* bezeichnete Wirkung von Kunst auf die menschliche Wahrnehmung ist das Schlüsselkonzept der strukturalistischen Definition von Ästhetik, wie in den folgenden Betrachtungen besprochen werden soll.

Seit den 1930er Jahren befassen sich auch Neurowissenschaftler mit der Frage nach den zerebralen Prozessen, die mit ästhetischen Erfahrungen einhergehen (z.B. Buswell, 1935; Wooding et al, 2002; und Yarbus, 1967). Interessanterweise identifizierten auch sie das Potenzial von Kunst, die menschliche Aufmerksamkeit zu bannen, als wesentliches Merkmal der ästhetischen Erfahrung. Während die Strukturalisten von der bereits erwähnten *Verfremdung* sprachen, ist hier von *inherent interestingness* (Hochberg, 1978) die Rede. Inzwischen hat sich das Forschungsgebiet der Neuroästhetik als eine Disziplin etabliert, die einerseits versucht, die zerebralen Prozesse ästhetischer Erfahrung zu ergründen (Zeki, 2001). Andererseits soll untersucht werden, durch welche Reizkonstellationen eine ästhetische Wahrnehmung bedingt ist (Graham und Redies, 2010). In dieser Fragestellung besteht der Fokus der vorliegenden Arbeit: Mittels Fourier-Analyse wurde der Versuch unternommen, jene strukturalen Eigenschaften visueller Stimuli zu identifizieren, die das zu Sehende für das menschliche Gehirn ‚interessant‘ (Hochberg, 1978) bzw. im positiven Sinne ‚fremd‘ (Jakobson, 1992) erscheinen lassen könnten.



Als wegbereitend für das Forschungsfeld des Strukturalismus erwies sich um 1900 die *Semiotik* oder *Zeichenlehre*, die keine eigenständige akademische Disziplin darstellt, jedoch als Teilgebiet der Erkenntnistheorie sowie der Sprachphilosophie und Sprachwissenschaft angesehen wird. Darüber hinaus haben ihre Ansätze zu Beginn des 20. Jahrhunderts die Entwicklung philosophischer Schulen wie jene des Prager Strukturalismus maßgeblich beeinflusst. Die Schriften des amerikanischen Philosophen und Mathematikers Charles Sanders Peirce, *Theory of signs* sowie *Theory of meaning* (Oehler, 1993), gelten dabei als theoretische Grundlage für die Werke der Hauptvertreter der Semiotik, Ferdinand de Saussures (Wunderli, 2013) und Roland Barthes (1979).

Die Semiotik untersucht sämtliche Zeichensysteme, die das Leben des Menschen prägen; nicht weniger relevant ist daher die Erforschung der menschlichen Wahrnehmung und Kommunikation. Die erwähnten Zeichensysteme beinhalten neben Verkehrsschildern oder Schriftsystemen alle Arten akustischer, visueller, gestischer und sprachlicher Zeichen (Wunderli, 2013). Gemäß der Theorie des Schweizer Sprachwissenschaftlers Ferdinand de Saussure verfügt jedes Zeichen unabhängig von seinem jeweiligen Medium oder seinen physikalischen Eigenschaften über denselben Aufbau. Das Zeichen in seiner Gesamtheit bezeichnet Saussure als *Sème*. Dieses gliedert sich einerseits in die mediale Seite, das sogenannte *Signifikant* oder *Bezeichnende*. Dabei handelt es sich um die akustische, visuelle oder haptische Beschaffenheit des Zeichens, welche durch das menschliche Sinnessystem wahrgenommen wird. Andererseits verfügt das Zeichen über eine mentale Seite, das *Signifikat* oder *Bezeichnete*, das imaginäre Bild also, welches durch den vorherigen Sinneseindruck evoziert wird. Der Bezug zwischen Signifikant und Signifikat besteht nicht naturgemäß, sondern wurde *willkürlich* durch die Individuen eines Zeichenkollektivs bestimmt (Wunderli, 2013). So evoziert beispielsweise das sprachliche Zeichen ‚Baum‘ – unabhängig davon, ob es sich um einen lautlichen oder visuellen Code oder um ein Piktogramm handelt – im Geist des Individuums das imaginäre Bild eines Baums. Wichtig ist, sich hierbei zu vergegenwärtigen, dass das sprachliche Zeichen nicht ‚den Gegenstand an sich‘ darstellt, sondern dass es lediglich den symbolischen Verweis auf diesen Gegenstand im Sinne einer mentalen, bildlichen Vorstellung evoziert. Saussure selbst hat diese zentrale These in seiner Vorlesungsreihe *Cours de linguistique générale* (Grundfragen der allgemeinen Sprachwissenschaft) wie folgt formuliert: „[Die Sprache ist] eine Form, keine Substanz“ (Saussure, 2001). Die folgenden Abbildungen sollen diesen Bezug verdeutlichen:



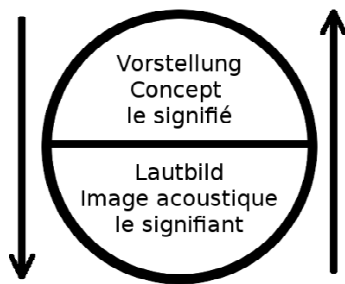


Abb. 4 - Komponenten des Zeichens (Wikimedia Commons, <http://commons.wikimedia.org/wiki/File:Signifikat.png>, 15.08.2013)

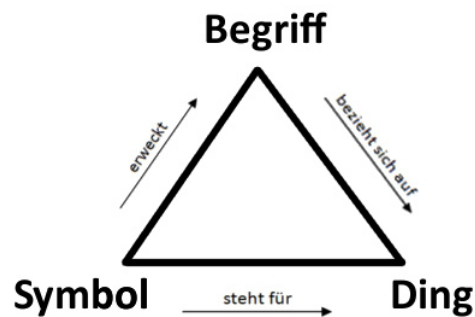


Abb. 5 - Das semiotische Dreieck (Wikimedia Commons, <http://commons.wikimedia.org/wiki/File:Semiotischesdreieck.jpg>, 15.08.2013)

Wie eingangs erwähnt, geht die Lehre des Strukturalismus davon aus, dass der Mensch in seinem täglichen Leben in ein Zeichensystem eingebunden ist, welches eine spezifische Struktur bzw. Komposition von Zeichen aufweist. Dem gegenüber stehen jene Zeichensysteme, die sich zwar derselben physikalischen Codes bedienen, also beispielsweise akustische oder visuelle Zeichen enthalten, deren Komposition oder Struktur sich aber von dem alltäglichen System unterscheidet. Sie werden durch das Individuum zwar erkannt und lassen sich zum Teil mit dem Gewohnten in Einklang bringen, sie erscheinen jedoch letztlich ‚anders‘ im Sinne einer ‚Verfremdung‘ des Alltäglichen. Der Strukturalismus geht davon aus, dass dieser *Verfremdungseffekt* die notwendige Voraussetzung einer ästhetischen Erfahrung darstellt (Holdheim, 1974).

Aus diesen Überlegungen lassen sich verschiedene Funktionen von Zeichen ableiten, wobei die Definition des Begriffs ‚Funktion‘ jener der Mathematik entspricht. Eine Funktion ist demnach die Beziehung zwischen zwei Variablen – dem Signifikant und dem Signifikat. In Anlehnung an das Kommunikationsmodell Bühlers (1999) unterscheidet die Sprachtheorie drei Zeichenfunktionen. Zum ersten kann das sprachliche Zeichen reine *Darstellungsfunktion* erfüllen, was Bühler selbst wie folgt formulierte: „Die Sprache bezieht sich auf die Welt, auf die Gegenstände und Sachverhalte. Dadurch erfüllt sie ihre Darstellungsfunktion“ (Kocsány, 2010). Ein Signifikat evoziert also direkt ein mentales Bild. Diese *referentielle Funktion* von Sprache hat demnach *symbolischen* Charakter. Die *expressive* bzw. *emotive Funktion* des sprachlichen Zeichens dient dem Sender zur Selbstkundgabe, was bewusst oder unbewusst geschehen kann. Die *Appellfunktion* des sprachlichen Zeichens kommt zum Tragen, sobald der

Sender versucht, den Empfänger in einer bestimmten Weise zu beeinflussen, beispielsweise im Sinne eines Befehls oder einer Bitte (Kocsány, 2010).

Im Zusammenhang mit Kunst führte der tschechoslowakische Literaturwissenschaftler Jan Mukařovský, seinerseits ein Vertreter des Prager Strukturalismus, die *ästhetische Funktion* ein. Für ihn stellt ein Kunstwerk ein so genanntes *Werk-Zeichen* dar. Die ästhetische Funktion des Werk-Zeichens gliedert sich in zwei Aspekte: *Autonomie* und *Kommunikation*. Aufgrund des autonomen Aspekts agiert das Kunstwerk "als Mittler zwischen den Mitgliedern des gleichen Kollektivs[...]" (Mukařovský, 1977). Bestimmte Darstellungen bzw. Werk-Zeichen sind im kollektiven Gedächtnis einer Gruppe von Individuen verankert und lösen dadurch bestimmte Emotionen und Assoziationen in den Betrachtern aus. Dieser Effekt bedingt nicht zuletzt das Gemeinschaftsgefühl der Gruppe. Der kommunikative Aspekt der ästhetischen Funktion beschreibt die Fähigkeit von Kunst, Bezug auf die Realität zu nehmen, Tendenzen aufzuzeigen und davon ausgehend Utopien oder Dystopien zu entwerfen. Mukařovský spricht hier von einer "unbestimmten Realität, auf die das Kunstwerk hinweist" und er bezieht sich dabei auf den "Gesamtkontext der sogenannten sozialen Erscheinungen: z.B. Philosophie, Politik, Religion, Wirtschaft usw." (Mukařovský, 1977).

Die Experimente der vorliegenden Arbeit konzentrierten sich vor allem auf den autonomen Aspekt der ästhetischen Funktion von Kunstwerken bzw. Werk-Zeichen und damit auf die Vermittlerfunktion der Zeichen innerhalb eines Kollektivs. Konkret wird dieser Fokus durch die in der Einleitung formulierte Frage danach verdeutlicht, ob interkulturelle Unterschiede bei der Bewertung von Zeichensystemen mit künstlerischem Anspruch bestehen.

Das zweite Ziel bestand im experimentellen Nachweis der strukturalen Unterschiede zwischen Zeichensystemen des alltäglichen Lebens ohne ästhetische Funktion und Zeichensystemen mit ästhetischem Anspruch. Methodisch wurde der Versuch unternommen, diese Unterschiede anhand der Fourier-Analyse nachzuweisen. Das folgende Kapitel bietet eine Zusammenschau und Diskussion der Ergebnisse.

## 2. Kunst versus Schrift: Alltägliche Druckschrift ist nicht skalierungsinvariant

Erwartungsgemäß unterscheidet sich das Fourier-Spektrum von Kunstbildern von jenem regulärer Druckschrift, wie in der Einleitung postuliert wurde. Die vorliegende Studie belegt diese Annahme systematisch und quantifiziert die gezeigten Unterschiede mithilfe eines Algorithmus, mit dem ein Computer bestimmte statistische Eigenschaften von Bildern erfassen kann. Wie in früheren Arbeiten (Graham und Field 2007; Redies et al, 2007 a, b; Graham und Redies, 2010) bereits gezeigt wurde, gehorcht die Abnahme der radiär gemittelten Signalstärke ( $1d$ ) der räumlichen Frequenzen in Kunstbildern mit zunehmender Ortsfrequenz dem Potenzgesetz  $1/f^p$ . Die Steigung  $[p]$  in einem doppelt logarithmischen Koordinatensystem beträgt etwa  $-2$ . Daraus folgt, dass das Fourier-Spektrum dieser Bildkategorie skalierungsinvariant ist. Kunstbilder teilen diese Eigenschaft mit anderen ästhetisch ansprechenden Bildkategorien, wie beispielsweise *graphic novels* und *comics* (Koch et al, 2010), aber auch mit Photographien komplexer natürlicher Szenen. Darüber hinaus wurden skalierungs-invariante oder *fraktale* Strukturen auch bei anderen künstlerischen Ausdrucksformen wie der Musik (Voss und Clarke, 1975; Beauvois, 2007), der Architektur (Joye, 2007) sowie bei der amerikanischen Gebärdensprache (Bosworth et al, 2006) nachgewiesen.

Die wissenschaftlichen Bemühungen, skalierungs-invariante Strukturen in verschiedenen Kunstformen nachzuweisen, steht in der Denktradition des irisch-britischen Philosophen und Schriftstellers Edmund Burke (1757), der im 18. Jahrhundert eine biologische Basis für die ästhetische Urteilskraft des Menschen postulierte. Als weiteren Vertreter dieses Diskurses gilt es für das 19. Jahrhundert den Begründer der experimentellen Ästhetik, Gustav Theodor Fechner, zu nennen. Im Rahmen seiner Versuche, Gesetzmäßigkeiten des ästhetischen Empfindens zu definieren, nahm Fechner bereits die Idee der Verfremdung, die das Ästhetische ‚inhärent interessant‘ erscheinen lässt, vorweg, indem er 1876 in seinem Werk *Vorschule der Ästhetik* schlussfolgert : „Etwas muss sowohl von der Stärke als auch von der Qualität her aufmerksamkeitswürdig sein, damit ich mich ihm zuwende.“ (Lennig, 1994). In den 1980er Jahren erweiterte der deutsche Philosoph Gregor Paul diese Idee um die Annahme, dass eine ästhetische Erfahrung nicht nur durch spezifische biologische Stimuli hervorgerufen wird, sondern ihrem Wesen nach als universell zu betrachten ist und interkulturell den gleichen Gesetzmäßigkeiten gehorcht (Paul, 1985). In den 1990er Jahren

schlossen sich verschiedene Vertreter der Neurowissenschaften diesem Diskurs an und argumentierten auf Basis naturwissenschaftlicher Überlegungen, dass das Zusammenspiel zwischen der menschlichen Wahrnehmung und den dadurch ausgelösten, neurologischen Prozessen das biologische Korrelat der ästhetischen Erfahrung darstellen muss (Livingstone, 1988; Gregory et al, 1995; Werner und Ratliff, 1999; Zeki, 1999 a, b; Livingstone, 2002; Redies, 2007).

Der Aspekt der menschlichen Wahrnehmung als entscheidendes Schlüsselkonzept zur Erforschung der ästhetischen Urteilskraft warf unweigerlich die Frage nach der Beschaffenheit jener Stimuli auf, die in der Lage sind, die zerebralen Prozesse zur Erfahrung von Ästhetik auszulösen. Bereits seit dem 18. Jahrhundert besteht ein Konsens zwischen Künstlern und Philosophen, dass es sich bei derartigen visuellen Stimuli um einzelne Bildelemente handelt, die sich untereinander zu einer harmonischen Gesamtkomposition zusammenschließen (Burke, 1757; Nietzsche, 1888; Matisse, 1908; Kandinsky, 1914; Paul, 1988). Dieses Postulat konnte durch das später in der Psychologie definierte *Gestalt-Gesetz* untermauert werden und lässt sich nicht zuletzt auf natürliche Szenen anwenden: Die Gestaltpsychologie geht davon aus, dass die menschliche Wahrnehmung Sinneseindrücke nicht wahllos erfasst, sondern sie unmittelbar zueinander in Beziehung stellt, um sie dadurch in eine bestimmte Ordnung – oder *Gestalt* – zu integrieren. So werden vertikale und horizontale Linien eher in die räumlichen Gestalt eines Quaders integriert, als sie zusammenhangslos ‚nebeneinander‘ oder ‚nacheinander‘ wahrzunehmen. Das Gehirn ist demnach stets bestrebt, Fragmente in einen stimmigen Gesamtkontext einzubetten. Kunstbilder werden diesem Anspruch offenbar in hohem Maße gerecht und teilen diese Eigenschaft mit natürlichen Szenen. Auch bei dieser Bildkategorie sind die einzelnen Bildelemente nicht zufällig angeordnet, sondern in einen Gesamtkontext eingebettet (Sigman et al, 2001).

Der Gesamtkontext zwischen den einzelnen Bildelementen entspricht der Komposition bzw. der globalen Struktur des Bildes. Diese globalen Strukturen werden in der Informatik als Strukturen der *zweiten Ordnung* oder als Strukturen *höherer Ordnung* bezeichnet. Sie stehen den Strukturen der *ersten Ordnung* gegenüber, bei denen es sich um fragmentarische (lokale) Bildelemente handelt. Die Fourier-Analyse hat sich als ein geeignetes Verfahren herausgestellt, um die in der Bildinformation enthaltenen Strukturen verschiedener Ordnung

zu quantifizieren. Die Fourier-Analyse natürlicher Szenen ergab erstens eine Dominanz kardinaler Orientierungen entlang des Spektrums der untersuchten Bilder (Burton und Moorhead, 1987; Field, 1987; Tolhurst et al, 1992; Ruderman und Bialek, 1994). Als eine weitere, signifikante Eigenschaft stellte sich zweitens heraus, dass die Signalstärke (*power*) mit dem Quadrat der Ortsfrequenz  $f$  abnimmt, dabei also dem Potenzgesetz  $1/f^2$  gehorcht, selbst wenn zwischen Einzelbildern eine gewisse Varianz verzeichnet wurde (Olshausen und Field, 1996, 2000; Simoncelli und Olshausen, 2001). Diese statistische Besonderheit korreliert mit dem hohen Anteil der in der Bildinformation enthaltenen, skalierungs-invarianten oder fraktalen Strukturen, die Strukturen höherer Ordnung darstellen. Auf Ebene der unmittelbaren Bildinformation lässt sich Fraktalität dahingehend interpretieren, dass das Verhältnis von hohen Raumfrequenzen (d.h., der feinen Bildstrukturen) zu tiefen Raumfrequenzen (d.h., der groben Bildstrukturen) in den einzelnen Bildausschnitten denen der Gesamtkomposition ähnlich ist. Es spielt also insofern keine Rolle, ob das Bild als Ganzes oder in Teilen betrachtet wird, da das Fourier-Frequenzspektrum konstant bleibt und somit skalierungsinvariant ist.

Die Ergebnisse zur Erforschung der statistischen Eigenschaften von Kunstbildern lassen sich wie folgt zusammenfassen: In verschiedenen Studien der letzten 10 Jahre konnte gezeigt werden, dass der Abfall der Signalstärke (*Power*) des Fourier-Spektrums verschiedener Kategorien von Kunstbildern dem Potenzgesetz  $1/f^2$  (Graham und Field, 2007; Redies et al 2007) gehorcht, wobei der kulturelle Hintergrund der analysierten Bilder sowie die verwendete Technik nur einen vergleichsweise geringen Einfluss auf die Ergebnisse ausüben. So fällt bei Kunstbildern westlicher und fernöstlicher Herkunft ebenso wie bei natürlichen Szenen die Signalstärke im Mittel mit einer Steigung von etwa -2 ab. Dieselbe Steigung wurde für die Darstellungen japanischer Mangas sowie für *comics* westlichen Ursprungs ermittelt. Bei beiden Bildkategorien handelt es sich um Genres, die im Betrachter neben einer ästhetischen Erfahrung zusätzlich positive Gefühle hervorrufen sollen (Koch et al, 2010).

Im Gegensatz zu Kunst zeigt sich bei regulärer Druckschrift ein steilerer Kurvenverlauf im hohen Frequenzbereich des Signalspektrums mit Werten für die Steigung von um -3,5, während die Steigung im niedrigen Frequenzbereich mit Werten von durchschnittlich -1,2 flacher abfällt (siehe Tabelle 1, Abbildung 4 der Publikation). Die Steigungswerte der

horizontalen, vertikalen und schrägen Orientierungen unterscheiden sich nur geringfügig. Ähnliche Ergebnisse für alltägliche Handschriften legen nahe, dass sich das Fourier-Spektrum von Gebrauchsschriften im niedrigen Frequenzbereich allgemein dahingehend von Kunstbildern unterscheidet, dass die Signalstärke schwächer ausfällt, es jedoch einen markanten Signalanstieg bei etwa 8 Perioden pro Bild gibt, der bei den vertikalen Orientierung auftritt und der Ortsfrequenz des Zeilenabstands entspricht.

Wie bereits erwähnt, werden im niedrigen Frequenzbereich feine Strukturen repräsentiert, während sich grobe, flächige Strukturen im hohen Frequenzbereich niederschlagen. Die unterschiedliche Signalstärke in diesen Bereichen kann dahingehend interpretiert werden, dass sich Gebrauchstexte im Vergleich zu den analysierten Kunstbildern weniger durch globale Strukturen auszeichnen, sondern fragmentarisch komponiert sind. Die Dominanz globaler Strukturen in Kunstbildern könnte dem physikalischen Korrelat der künstlerischen Komposition entsprechen, durch welche die einzelnen Bildelemente, wie in der Einleitung diskutiert, in einen harmonischen Gesamtzusammenhang einbettet werden. Dies wiederum könnte möglicherweise die Bildinformation für das menschliche Sehsystem *interessant* erscheinen lassen.

In einigen früheren Studien wurde die Meinung vertreten, dass sich alle Bilder, die durch Menschen geschaffen wurden, durch skalierungs-invariante Eigenschaften auszeichnen, da sich dies auf die natürlichen Bewegungsmuster der menschlichen Hand zurückführen ließe (siehe z.B. Graham und Field, 2008). Die Ergebnisse der vorliegenden Studie zeigen jedoch, dass nicht alle von Menschenhand geschaffenen Bilder oder ‚Textbilder‘ fraktale Strukturen in ihrem Fourier-Spektrum aufweisen. Dass handschriftliche Textbilder weniger globale Strukturen beinhalten ist nicht überraschend, zumal globale Strukturen die Funktionalität von Gebrauchsschrift nicht unterstützen; gelesen wird, wie einleitend geschildert, Wort für Wort im Sinne eines kontinuierlichen, linearen Fixationsvorgangs. Darüber hinaus ist auch der hohe Frequenzbereich des Fourier-Spektrums von Gebrauchstexten nicht skalierungs-invariant. Daraus lässt sich ableiten, dass von Menschen produzierte Bilder entsprechend ihres Verwendungszweckes skalierungs-invariant sein können, oder nicht. Interessanterweise wurde in früheren Studien gezeigt, dass die Manipulation des Fourier-Spektrums von Kunstbildern beim Betrachter Unbehagen hervorrufen kann, sobald die Abweichung von der ursprünglichen Fraktalität des Bildes einen bestimmten Schwellenwert

übersteigt (Fernandez und Wilkins, 2008; Juricevic et al, 2010; O'Hare und Hibbard, 2011): Die Betrachtung von Bildern wurde als visuell unbehaglich beschrieben, sobald das manipulierte Spektrum einen gekrümmten Kurvenverlauf sowie eine Zunahme der Signalstärke um 3 bis 4 Perioden pro Grad Seewinkel aufwies (Fernandez und Wilkins, 2008). Das Ausmaß des visuellen Unbehagens wird in diese Fällen mit einer vermeintlich mangelhaften, künstlerischen Leistung in Zusammenhang gebracht (Fernandez und Wilkins, 2008).

Auch in der vorliegenden Studie wurden gekrümmte Kurvenverläufe an bestimmten Bildkategorien beobachtet, zum Beispiel an Gebrauchstexten, Kalligraphie und ornamentaler Kunst. Allerdings rufen diese Bildkategorien tendenziell kein visuelles Unbehagen hervor. Es bleibt daher fraglich, ob die Ergebnisse dieser Studie mit den Ergebnissen von Fernandez und Wilkins (2008) in Bezug gestellt werden können. Dennoch erlaubt die Zusammenschau der Ergebnisse beider Untersuchungen den Schluss, dass sich Bildkategorien ohne skalierungs-invariante Eigenschaften im Allgemeinen durch einen niedrigen ästhetischen Gehalt auszeichnen.

Der Umkehrschluss, Fraktalität korreliere direkt mit Ästhetik, erweist sich jedoch nicht als haltbar, zumal Darstellungen mit einem skalierungs-invarianten Fourier-Spektrum nicht zwingend als ästhetisch empfunden werden, wie Lee und Mumford (1999) anhand computergenerierter Bilder zeigten. Dieser Umstand führte zu der Hypothese, dass Skalierungsinvarianz die Begleiterscheinung einer bislang noch unerforschten Eigenschaft von Bildern mit künstlerischen Anspruch sein könnte, die im Betrachter eine ästhetische Erfahrung hervorruft (Redies et al, 2007a).

### **3. Steigungsdifferenz und Isotropie: Korrelate künstlerischen Anspruchs**

Die Betrachtung von Text erwies sich als vielversprechend im Hinblick auf die Erforschung ästhetischer Wahrnehmung, zumal sich herausstellte, dass es einen fließenden Übergang gibt zwischen Gebrauchsschriften und ästhetischen Schriften mit Ornamenten oder Kalligrafie bis hin zu Kunstbildern. Meine Ergebnisse zeigen (siehe Tabelle 1 und Abbildungen 3, 4), dass die statistischen Eigenschaften von Textbildern mit zunehmendem ästhetischen Anspruch sich den Eigenschaften von Kunstbildern annähern. Um diese Annäherung zu quantifizieren, wurde die Differenz der Steigung im niedrigen und hohen Frequenzbereich

des radiär gemittelten Signalspektrums bestimmt. Dabei stellte sich heraus, dass sich die Steigungsdifferenz in allen Orientierungen mit zunehmendem künstlerischem Anspruch dem Zahlenwert 0 annähert (siehe Tabelle 1 und Abbildung 4). Dieser Wert entspricht einer geraden Linie und verweist auf den annähernd linearen Abfall der Signalstärke gemäß  $1/f^2$  über das gesamte Spektrum, wie bereits für natürliche Szenen beschrieben.

Ein ähnlicher Übergang konnte auch an horizontalen und vertikalen Orientierungen, die ihrerseits ein Maß für die Anisotropie eines Bildes darstellen, beobachtet werden (siehe Abbildung 3). Bei Textbildern geht die Differenz der Steigung zwischen vertikalen und horizontalen Orientierungen wahrscheinlich auf die Periodizität der Zeilen und Buchstaben (für die vertikale Orientierung) zurück sowie auf deren Strichbreite und Zeichenabstände (für die horizontale Orientierung). Im Gegensatz zu Gebrauchstext ist die Steigung des Fourier-Spektrums in aller räumlichen Orientierungen bei Kunstbildern annähernd gleich stark, wie bereits von Koch et al (2010) gezeigt, der Kunstwerke mit anderen Bildkategorien verglich. Diese Beobachtung ist keineswegs trivial, zumal es Künstlern nicht schwerfallen dürfte, Bilder zu schaffen, in denen lediglich eine räumliche Orientierung dominiert. Ob Anisotropie jedoch als notwendige Voraussetzung für eine ästhetische Erfahrung interpretiert werden muss, und falls ja, für welche Bildkategorien sie eine derartige Voraussetzung darstellen könnte, bleibt Gegenstand zukünftiger Studien.

Werke des Abstrakten Expressionismus wurden in die Analyse dieser Studie mit einbezogen, weil diese aufgrund ihrer spezifischen Komposition aus repetitiven Bildelementen der Struktur von Textbildern ähneln. Allerdings stellte sich heraus, dass die statistischen Eigenschaften von Bildern des Abstrakten Expressionismus mehr Ähnlichkeit mit Kunstbildern als mit Gebrauchsschriften aufweisen, was den Schluss nahelegt, dass ihre Komposition durch einen hohen Anteil globaler Strukturen geprägt ist. Dieses Ergebnis ist ebenfalls nicht trivial, wenn man bedenkt, dass man dieselben Bildelemente nicht zwingend im Sinne einer globalen Bildkomposition hätte anordnen müssen, sondern sie beispielsweise fragmentarisch nebeneinander stellen könnte. Damit stimmen die Ergebnisse meiner Studie mit der ursprünglich von Richard Taylors gemachten Beobachtung überein, dass Werke des Abstrakten Expressionismus, wie beispielsweise die Klecksbilder eines Jackson Pollock, durch fraktale Strukturen geprägt sind (Taylor, 2002; Mureika, 2005; Alvarez-Ramirez et al, 2008).



Ein weiteres Beispiel für eine Bildkategorie mit repetitiven, flächigen Strukturen sind ornamentale Verzierungen, wobei hier der eigentlich künstlerische Gehalt tendenziell hinter die Unterhaltungsfunktion der Darstellung tritt: Besonders die Malereien auf chinesischem Teegeschirr erwecken in vielen Fällen den Eindruck, eher Stoff für Tischgespräche liefern zu wollen, als den Anspruch auf eine hochrangige künstlerische Darstellung zu erheben. Die Steigungsdifferenz ornamentaler Kunst aus allen drei Kulturkreisen liegt zwischen den Werten von Kunstbildern und Gebrauchsschrift. Die Anisotropie im Sinne der Differenz zwischen horizontalen und vertikalen Orientierungen liegt näher bei den Ergebnissen für Kunst als für Gebrauchsschrift (siehe Abbildung 3). Ähnliche Zwischenwerte ergeben sich für die Kalligrafie, eine Schriftkategorie mit hohem künstlerischem Anspruch. Dabei bleibt es Gegenstand zukünftiger Forschung, ob es auch bei anderen Bildkategorien einen ähnlichen Zusammenhang zwischen den statistischen Gesetzmäßigkeiten und dem künstlerischem Anspruch gibt.

#### **4. Ästhetische Erfahrung: Kein Anhalt für kulturellen Einfluss**

In der vorliegenden Studie wurde eine horizontale Herangehensweise gewählt, um zeitgenössische westliche, arabische und fernöstliche Textbeispiele aus den verschiedenen Schriftkategorien zu vergleichen (siehe Tabelle 1 und Abbildungen 3,4). Die analysierten Schriftproben beinhalten sowohl schlichte als auch verzierte Schriftarten, wurden in Schriften mit und ohne Serifen eingeteilt und entstammen 68 verschiedenen Alphabeten. Im Zuge der vertikalen Annäherung an die Fragestellung dieser Studie wurden außerdem mittelalterliche Manuskripte, kalligrafische Werke und Handschriften aus der Reformationszeit mit einbezogen. Insgesamt wurden 11 Datensätze mit 1598 Textbildern analysiert, die jeweils zwischen 13 und 253 Schriftproben enthielten (siehe Tabelle 1).

Die Fourier-Analyse lieferte ähnliche Ergebnisse für die Schriftproben aller drei Kulturräume (siehe Tabelle 1 und Abbildungen 3,4). Besonders ähnlich erwiesen sich die Steigungsdifferenzen von lateinischer und internationaler Druckschrift mit Serifen sowie von gedruckten Zierschriften, Kalligrafie und ornamentaler Kunst aus jeweils allen Kulturräumen. Darüber hinaus belegen die Ergebnisse, dass das Fourier-Spektrum von Werken bildender Kunst aus dem fernen Osten sowie aus der westlichen Welt gleichermaßen durch skalierungs-invariante Strukturen geprägt ist (Redies et al, 2007b; Graham und Field, 2008).

Die Zusammenschau der Ergebnisse lässt sich dahingehend interpretieren, dass sowohl für das Lesen als auch für die ästhetische Bewertung von Schrift möglicherweise universelle Wahrnehmungsmechanismen existieren, die keine interkulturellen Unterschiede aufweisen. Für Schriftsysteme verschiedener kultureller Herkunft lässt sich des Weiteren verallgemeinern, dass ein Zeichen aus durchschnittlich drei Einzelstrichen besteht und dass es sich dabei um redundante Einzelelemente handelt. So konnten Changizi und Shimojo (2005) zeigen, dass Buchstaben durch das menschliche Sehsystem selbst dann noch erkannt werden, wenn die Hälfte der einzelnen Striche weggelassen wird. Verschiedene Spekulationen gehen davon aus, dass diese Redundanz dem Selektionsdruck geschuldet sein könnte, dem Schriftzeichen unterliegen: Buchstaben bestehen aus redundanten Einzelelementen, um erstens leicht durch das menschliche Sehsystem erkannt zu werden und zweitens dem motorischen System der Hand entgegenzukommen (Changizi und Shimojo, 2005; Changizi et al, 2006; Graham und Redies, 2010).

Nicht zuletzt sei erwähnt, dass die physikalische Beschaffenheit visueller Stimuli einen entscheidenden Faktor für die Modulation funktionaler Antworten in zerebralen Gehirngebieten darstellt. So reagieren zum Beispiel jene Areale des Gehirns, die als Lesezentrum angesehen werden können, besonders sensibel auf bestimmte Ortsfrequenzen und aktivieren unmittelbar den linken ventralen occipitotemporalen Cortex, wobei dieser bei der Erkennung von Wörtern wiederum unmittelbar mit dem unteren Frontallappen interagiert (Seghier und Price, 2011; Woodhead et al, 2011; Horie et al, 2012).

Die Untersuchung der statistischen Eigenschaften der in der vorliegenden Studie analysierten Bildkategorien leistet im neurologischen Sinne einen Beitrag zur Erforschung der psychophysischen Grundlagen jener zerebralen Prozesse, die als Reaktion auf geschriebene Worte und ähnliche Stimuli im menschlichen Gehirn ablaufen. Darüber hinaus kann in zukünftigen Untersuchungen gefragt werden, inwieweit die statistischen Eigenschaften der analysierten Bildkategorien mit jenen Strukturen korrelieren, die den von den Strukturalisten postulierten Verfremdungseffekt nicht-alltäglicher Stimuli erzeugen und die Sehinformation für das menschliche Gehirn interessant erscheinen lassen (siehe Kapitel 1 der Diskussion). Die Erlangung tieferer Einblicke in die entsprechenden neuropsychologischen Prozesse bleibt zukünftigen Studien vorbehalten

## LITERATURVERZEICHNIS

- Alvarez-Ramirez J, Ibarra-Valdez C, Rodriguez E, Dagdug L. 2008b. 1/f-noise structure in Pollock's drip paintings. *Physica*, 387: 281-295.
- Barthes R. 1979. *Elemente der Semiologie*. Frankfurt am Main: Sydikat.
- Bosworth RG, Bartlett MS, Dobkins KR. 2006. Image statistics of American Sign Language: comparison with faces and natural scenes. *Journal of the Optical Society of America*, 23: 2085-96.
- Bouma H. 2004. Interaction effects in parafoveal letter recognition. *Nature*, 226: 177-8.
- Bühler K. 1999. *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Stuttgart: Fischer.
- Burke E. 1757. *A philosophical enquiry into the origin of our ideas of the sublime and beautiful*. London: Dodsley.
- Burton GJ, Moorhead IR. 1987. Color and spatial structure in natural scenes. *Applied Physiology*, 26: 157-170.
- Buswell GT. 1935. *How People Look at Pictures*. Chicago: University of Chicago Press.
- Campbell FW, Robson JG. 1968. Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197: 551-566.
- Changizi MA, Shimojo S. 2005. Character complexity and redundancy in writing systems over human history. *Proceedings of the Royal Society B: Biological Sciences*, 272: 267-275.
- Changizi MA, Zhang Q, Ye H, Shimojo S. 2006. The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *The American Naturalist*, 167: E117-E139.
- Chung ST, Tjan BS. 2007. Shift in spatial scale in identifying crowded letters. *Vision Research*, 47: 437-451.
- Chung ST, Tjan BS. 2009. Spatial-frequency and contrast properties of reading in central and peripheral vision. *Journal of Vision*, 9: 16.1-16.19.
- Conway BR, Livingstone MS. 2007. Perspectives on science and art. *Current Opinion in Neurobiology*, 17: 476-82.
- Fernandez D, Wilkins AJ. 2008. Uncomfortable images in art and nature. *Perception*, 37: 1098-113.

- Field DJ. 1987. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A - Optics and Image Science*, 4: 2379-2394.
- Goldstein EB. 2002. *Wahrnehmungspsychologie*. Heidelberg: Spektrum.
- Goodale MA, Milner AD. 1992. Separate visual pathways for perception and action. *Trends in Neuroscience*, 15: 20-5.
- Graham DJ, Field DJ. 2007. Statistical regularities of art images and natural scenes: spectra, sparseness and nonlinearities. *Spatial Vision*, 21: 149-164.
- Graham DJ, Field DJ. 2008. Variations in intensity statistics for representational and abstract art, and for art from the Eastern and Western hemispheres. *Perception*, 37: 1341-1352.
- Graham D, Redies C. 2010. Statistical regularities in art: relations with visual coding and perception. *Vision Research*, 50: 1503-1509.
- Gregory RL, Harris J, Heard P, Rose D. 1995. *The Artful Eye*. Oxford: Oxford University Press.
- Hochberg J. 1978. Visual art and the structures of the mind. In: *The Arts, Cognition, and Basic Skills* (Madeja, Ed.). New York: Cambridge University Press. S. 151-172.
- Holdheim WW. 1974. The concept of poetic estrangement. *Comparative Literature Studies*, 11: 320-325.
- Horie S, Yamasaki T, Okamoto T, Kan S, Ogata K, Miyauchi S, et al. 2012. Distinct role of spatial frequency in dissociative reading of ideograms and phonograms: an fMRI study. *Neuroimage*, 63: 979-988.
- Hoyer PO, Hyvärinen A. 2002. Sparse coding of natural contours. *Neurocomputing*, 44-46: 459-466.
- Hubel DH, Wiesel TN. 1959. Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology*, 148: 574-591.
- Hubel DH, Wiesel TN. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160: 106-154.
- Jainta SW, Jaschinski W, Wilkins AJ. 2010. Periodic letter strokes within a word affect fixation disparity during reading. *Journal of Vision*, 10: 2.
- Jakobson R. 1992. *Semiotik: Ausgewählte Texte 1919-1982*. Frankfurt am Main: Suhrkamp.

- Joye Y. 2007. Architectural lessons from environmental psychology: the case of biophilic architecture. *Review of General Psychology*, 11: 305-328.
- Juricevic I, Land L, Wilkins A, Webster MA. 2010. Visual discomfort and natural image statistics. *Perception*, 39: 884-899.
- Kandel E. 2012. *Das Zeitalter der Erkenntnis*. München: Siedler.
- Kandinsky W. 1914. *The Art of Spiritual Harmony*. London: Constable.
- Koch M, Denzler J, Redies C. 2010. 1/f 2 Characteristics and isotropy in the fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs. *PLoS ONE*, 5: e12268.
- Kocsány P. 2010. *Grundkurs Linguistik: ein Arbeitsbuch für Anfänger*. Paderborn: Fink.
- Kuffler SW, Nicholls JG. 1966. The physiology of neuroglial cells. *Ergebnisse der Physiologie, Biologischen Chemie und Experimentellen Pharmakologie*, 57: 1-90.
- Lee AB, Mumford D. 1999. An occlusion model generating scale-invariant images. In: *Proceedings of the IEEE Workshop on Statistical and Computer Theories of Vision*, Fort Collins: 1-20.
- Lennig P. 1994. *Von der Metaphysik zur Psychophysik. Gustav Theodor Fechner (1801-1887)*. Bern: Verlag Peter Lang.
- Levine DN. 2007. Sherrington's "The Integrative action of the nervous system": A centennial appraisal. *Journal of the Neurological Sciences*, 253:1-6.
- Livingstone MS, Conway BR. 2004. Was Rembrandt stereoblind? *The New England Journal of Medicine*, 351: 1264-1265.
- Livingstone MS. 2002. *Vision and Art: The Biology of Seeing*. New York: Harry N. Abrams.
- Livingstone MS. 1988. Art, illusion and the visual system. *Scientific American*, 258: 78-85.
- Maffei L, Fiorentini A. 1973. The visual cortex as a spatial frequency analyzer. *Vision Research*, 13: 1255-1267.
- Majaj NJ, Pelli DG, Kurshan P, Palomares M. 2002. The role of spatial frequency channels in letter identification. *Vision Research*, 42: 1165-1184.
- Marr D. 1982. *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information*. New York: Freeman.
- Martinez-Conde S, Stephen LM, Hubel DH. 2004. The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, 5: 229-240.

- Matisse H. 1908. Notes d'un peintre. *La Grande Revue*, 2: 731-745.
- Milner AD, Goodale MA. 2008. Two visual systems re-viewed. *Neuropsychologia*, 46: 774-85.
- Mukařovský J. 1977. Studien zur strukturalistischen Ästhetik und Poetik. München: Ullstein.
- Mureika JR, Dyer CC, Cupchik GC. 2005. Multifractal structure in nonrepresentational art. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 72: 0461010-1-15.
- Nandy AS, Tjan BS. 2008. Efficient integration across spatial frequencies for letter identification in foveal and peripheral vision. *Journal of Vision*, 8: 3.1-3.20.
- Nietzsche F. 1888. Der Fall Wagner. Kritische Studienausgabe. Berlin: Deutscher Taschenbuchverlag.
- Oehler K. 1993. Charles Sanders Peirce. München: Beck.
- O'Hare L, Hibbard PB. 2011. Spatial frequency and visual discomfort. *Vision Research*, 51: 1767-1777.
- Olshausen BA, Field DJ. 1996. Natural image statistics and efficient coding. *Network*, 7: 333-339.
- Olshausen BA, Field DJ. 2000. Vision and the coding of natural images. *American Scientist*, 88: 238-245.
- Oruc I, Landy MS. 2009. Scale dependence and channel switching in letter identification. *Journal of Vision*, 9: 4.1-4.19.
- Ozawa S. 1994. 1/f fluctuation in Japanese brush calligraphy. *Research of Pattern Formation*. Tokyo: KTK Scientific Publishers, 571-579.
- Paul G. 1985. Der Mythos der modernen Kunst und die Frage nach der Beschaffenheit einer zeitgemäßen Ästhetik. Wiesbaden: Steiner.
- Paul G. 1988. Philosophical theories of beauty and scientific research on the brain. In: *Beauty and the Brain* (Epstein, Herzberger und Rentschler, Ed.). Basel: Birkhäuser. S. 15-27
- Paul G. 1998. Philosophische Ästhetik: Grundbegriffe und Grundpositionen. Stuttgart: Klett.
- Parraga CA, Troscianko T, Tolhurst DJ. 2000. The human visual system is optimised for processing the spatial information in natural visual images. *Current Biology*, 10: 35-38.
- Pelli DG, Burns CW, Farell B, Moore-Page DC. 2006. Feature detection and letter identification. *Vision Research*, 46: 4646-4674.



- Pritchard RM, Heron W, Hebb DO. 1960. Visual perception approached by the method of stabilized images. *Canadian Journal of Psychology*, 14: 67-77.
- Quiroga RQ, Pedreira C. 2011. How do we see art: an eye-tracker study. *Frontiers in Human Neuroscience*, 5: 98.
- Read H. 1954. Art and the evolution of consciousness. *The Journal of Aesthetics and Art Criticism*, 13: 143-155.
- Redies C. 2007. A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision*, 21: 97-117.
- Redies C, Hänisch J, Blickhan M, Denzler J. 2007a. Artists portray human faces with the Fourier statistics of complex natural scenes. *Network*, 18: 235-248.
- Redies C, Hasenstein J, Denzler J. 2007b. Fractal-like image statistics in visual art: similarity to natural scenes. *Spatial Vision*, 21: 137-148.
- Ruderman DL, Bialek W. 1994. Statistics of natural images - scaling in the woods. *Physical Review Letters*, 73: 814-817.
- Saussure F. 2001. *Grundfragen der allgemeinen Sprachwissenschaft*. Berlin: de Gruyter.
- Seghier ML, Price CJ. 2011. Explaining left lateralization for words in the ventral occipitotemporal cortex. *The Journal of Neuroscience*, 31: 14745-14753.
- Sigman M, Cecchi GA, Gilbert CD, Magnasco MO. 2001. On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences, USA* 98: 1935-1940.
- Simoncelli EP. 2003. Vision and the statistics of the visual environment. *Current Opinion in Neurobiology*, 13: 144-149.
- Solomon JA, Pelli DG. 1994. The visual filter mediating letter identification. *Nature*, 369: 395-397.
- Taylor RP. 2002. Order in Pollack's chaos: Computer analysis is helping to explain the appeal of Jackson Pollock's paintings. *Scientific American*, 287: 116-121.
- Tolhurst DJ, Tadmor Y, Chao T. 1992. Amplitude spectra of natural images. *Ophthalmic and Physiological Optics*, 12: 229-232.
- Tyler CW, Likova LT. 2007. Crowding: a neuroanalytic approach. *Journal of Vision*, 7: 16.1-16.9.
- Vinje WE, Gallant JL. 2000. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287: 1273-1276.

- Voss RF, Clarke J. 1975.  $1/f$  noise in music and speech. *Nature*, 258: 317-318.
- Werner JS, Ratliff F. 1999. Some origins of the lightness and darkness of colors in the visual arts and in the brain. *Techne*, 9-10: 61-73.
- Westheimer G, McKnee S. 1975. Visual acuity in the presence of retinal-image motion. *Journal of the Optical Society of America*, 65: 847-50.
- Wilkins AJ, Smith J, Willison CK, Beare T, Boyd A, Hardy G, et al. 2007. Stripes within words affect reading. *Perception*, 36: 1788-1803.
- Woodhead ZV, Wise RJ, Sereno M, Leech R. 2011. Dissociation of sensitivity to spatial frequency in word and face preferential areas of the fusiform gyrus. *Cerebral Cortex*, 21: 2307-2312.
- Wooding DS, Mugglestone MD, Purdy KJ, Gale AG. 2002. Eye movements of large populations: I. Implementation and performance of an autonomous public eye tracker. *Behavior Research Methods, Instruments & Computers*, 34: 509-517.
- Wunderli P. 2013. Ferdinand de Saussure: Cours de linguistique générale: Zweisprachige Ausgabe französisch-deutsch mit Einleitung, Anmerkungen und Kommentar. Tübingen: Narr.
- Yarbus AL. 1967. *Eye Movements and Vision*. New York: Plenum.
- Zeki S. 1999a. Art and the Brain. *Journal of Consciousness Studies*, 6-7: 76-96.
- Zeki S. 1999b. *Inner Vision*. Oxford: Oxford University Press.
- Zeki, S. 2001. Artistic Creativity and the Brain. *Science*, 293: 51-52.

## **Zitate**

- Seite 1: Read H. 1954. Art and the evolution of consciousness. *The Journal of Aesthetics and Art Criticism*, 13: 144.
- Seite 8: Zeki S. 2001. *Inner Vision: An Exploration of Art and the Brain*. New York: Oxford University Press, 113.

# LEBENS LAUF

Name: Tamara Melmer  
Geburtsdatum: 18.02.1984  
Geburtsort: Rum/ Österreich  
Staatsangehörigkeit: Österreich

## Schule

Juli 2002 Matura am Wirtschaftskundlichen Realgymnasium der Ursulinen in Innsbruck

## Studium

Oktober 2002 Aufnahme des Studiums Anglistik/ Amerikanistik sowie Slawistik/ Russisch an der Leopold-Franzens-Universität Innsbruck

April 2003 Aufnahme des Studiums der Komparatistik an der Leopold-Franzens-Universität Innsbruck

Dezember 2007 Diplomprüfung Anglistik/ Amerikanistik mit Erlangung des akademischen Grades der Magistra philosophiae

Januar 2008 Diplomprüfung Slawistik/ Russisch mit Erlangung des akademischen Grades der Bakkalaurea philosophiae

April 2008 Diplomprüfung Komparatistik mit Erlangung des akademischen Grades der Magistra philosophiae

Oktober 2008 Aufnahme des Studiums der Humanmedizin an der Friedrich-Schiller-Universität Jena

Juni 2015 Staatsexamen Humanmedizin an der Friedrich Schiller-Universität Jena

## Praktika und Auslandsaufenthalte

Februar - Mai 2004 Sprachstudium an der Staatlichen Universität Sankt Petersburg

Juli 2006 Sprachstudium an der Moskauer Akademie für Wirtschaftliche Zusammenarbeit

August - Oktober 2010 Praktikum bei der UNO/ UNODC in Wien im Rahmen des TRATNET-Projekts

August - Oktober 2012 Auslandsfamulatur in Indien im Rahmen des CALCUTTA VILLAGE-Projekts

August - Oktober 2013 Praktikum bei der WHO in Genf in der Abteilung für Arbeitsmedizin sowie im Rahmen des MANAGERIAL REFORM-Projekts

Jena, 10.06.2015: \_\_\_\_\_

## **EHRENWÖRTLICHE ERKLÄRUNG**

Hiermit erkläre ich, dass ich unter Berücksichtigung der Promotionsordnung der Medizinischen Fakultät der Friedrich-Schiller-Universität Jena die vorliegende Dissertation selbst angefertigt und alle benutzten Quellen als solche aufgelistet habe.

Folgende Mitglieder der Arbeitsgruppe Neuroästhetik haben mich bei der Planung und Durchführung der Experimente, sowie bei der Erstellung des Manuskripts unterstützt: Prof. Dr. Dr. Christoph Redies, Prof. Dr. Joachim Denzler, Ali Seyed Amirshahi, Julia Braun und Michael Koch.

Es wurde weder die Hilfe eines Promotionsberaters in Anspruch genommen, noch habe ich Dritte für Arbeiten bezahlt, die im Zusammenhang mit der vorgelegten Doktorarbeit stehen. Abschließend erkläre ich, dass ich diese Dissertation oder eine in wesentlichen Teilen ähnliche Arbeit bisher an keiner anderen Hochschule eingereicht habe.

Jena, 10.06.2015 \_\_\_\_\_